DOCUMENT RESUME

ED 311 537 EA 021 299

AUTHOR Ondrack, D. A.; Oliver, C.
TITLE A Review and Analysis of Performance Appraisal Processes, Volume III. Performance Appraisal for Professional Service Employees: Non-Technical Report. Professionalism in Schools Series.
INSTITUTION Ontario Dept. of Education, Toronto.
REPORT NO ISBN-0-7729-1090-1
PUB DATE 86
NOTE 84p.; For volumes 1 and 2, see EA 021 297-298.
AVAILABLE FROM Publications Sales, The Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, Ontario, Canada M5S 1V6.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Decentralization; Elementary Secondary Education; *Evaluation Criteria; *Evaluation Methods; Evaluation Utilization; Foreign Countries; *Formative Evaluation; Participative Decision Making; *Personnel Evaluation; *School Personnel; *Summative Evaluation
IDENTIFIERS *Ontario

ABSTRACT

The third of three volumes, this report summarizes the findings of, first, a review and analysis of published literature on performance appraisal in general and particularly on the use of appraisals in public education systems, and, second, a series of field-site investigations of performance appraisal systems in action. The field site studies of performance appraisal in action at 10 professional service organizations in Ontario and 7 public school systems in the United States mirror many of the results from the literature review. Three general conclusions are derived from the results of this study: (1) different approaches and priorities in appraisal systems are used for relevance to the local situation of a particular organization or school jurisdiction; (2) appraisal systems can be expected to go through stages of evolution as all participants have to go through a learning curve of experience with a system, and therefore, no one design can be expected to be correct at the outset or remain in force indefinitely; (3) an appraisal system developed jointly between supervisors and subordinates has a better chance of incorporating diverse but relevant points of view than a system developed centrally by top management alone. The greater the opportunity for participation by parties affected by a decision, the greater the potential quality and acceptability of the decisions. Because no system of appraisals is workable unless it is acceptable and legitimate to all staff, an appraisal system should be centrally developed with universal standards that are adaptable to change based on local operational specifics and requisites. (JAM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

EA 021 299

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Ministry
of
Education

A REVIEW AND ANALYSIS OF PERFORMANCE APPRAISAL PROCESSES VOLUME III

Performance Appraisal for Professional Service Employees: Non-Technical Report

Professionalism in Schools Series

D. A. ONDRACK, Principal Investigator

C. OLIVER, Research Associate

This research project was funded under contract by the Ministry of Education, Ontario.

This study reflects the views of the authors and not necessarily those of the Ministry of Education.

The Honourable Sean Conway, Minister

© QUEEN'S PRINTER FOR ONTARIO, 1986

Order Information:

Publications Sales
The Ontario Institute
for Studies in Education
252 Bloor Street West
Toronto, Ontario M5S 1V6
Will invoice on orders over $30.00.
Other orders must be accompanied by a
cheque or money order payable to
O.I.S.E.

Publications Services
880 Bay Street, 5th Floor
Toronto, Ontario M7A 1N8

(416) 965-6015
(Toll Free) 1-800-268-7540
(Toll Free from area code 807)
Ask operator for Zenith 67200
Order must be accompanied by a cheque
or money order payable to the
Treasurer of Ontario

ONO 4189

Canadian Cataloguing in Publication Data

Ondrack, D. A.
A review and analysis of performance appraisal processes

Includes bibliographies.
Partial contents: v. 2. Performance appraisal for professional service employees, field study report -- v. 3. Performance appraisal for professional service employees, non-technical report.
ISBN 0-7729-1087-1 (3 v. set). -- 0-7729-1088-X (v. 1). -- 0-7729-1089-8 (v. 2). -- 0-7729-1090-1 (v. 3)

1. Teachers--Rating of. 2. Employees, Rating of. I. Oliver, Christine Tomaszuk. II. Ontario. Ministry of Education. III. Title. IV. Title: Performance appraisal for professional service employees.

LB2838.052 1986 371.1'44 C86-099627-1

TABLE OF CONTENTS

	Page No.
Acknowledgements	vi
Abstract	vii
1. Introduction	1
2. Literature Review: A Summary	6
2.1 Purposes of Appraisal	6
2.2 Components of Appraisal	9
2.2.1 Type of Criteria	9
2.2.2 Method of Measurement	10
2.2.3 Sources of Data Measurement	13
2.3 Process of Appraisal	13
2.3.1 Choice of Appraisal	13
2.3.2 Appraisal Instruments	14
2.3.3 Appraisal Interviews	17
2.3.4 Appraisal Training	17
2.3.5 Assessment Period and Timing	18
2.4 Outcomes of Appraisal	19
2.5 Link to Administrator Evaluation	20
2.6 Comparison and Conclusions	22
3. Field Site Results	31
3.1 Introduction	31
3.1.1 The Business Group	31
3.1.2 The Colleges Group	32
3.1.3 The U.S. Schools Group	32
3.2 Procedure	33
3.3 Results	34
3.3.1 Initial Design and Implementation of the System	34
3.3.2 Current Objectives and Purposes of Appraisal	36
3.3.3 Appraisal Design and Techniques	36
3.3.4 Criteria or Standards for Appraisal	40
3.3.5 Sources of Data on Performance	42
3.3.6 Training and Preparation for Appraisal	45
3.3.7 Special Cases and Organization Justice	48
3.3.8 Evaluation and Maintenance of the Appraisal System	52
3.3.9 Trends and Outlook for the Apraisal Systems	53

Page No.

4. Conclusions 53

4.1 Objectives and Purposes of an Appraisal System 55

4.1.1 Situational Factors 55

4.1.2 Sense of Mission and Normative Models of Professional Service 56

4.1.3 Career Stages of Professionals 57

4.1.4 Absence of Mission or Normative Model 58

4.2 Design, Implementation, and Support of Appraisal Systems 58

4.3 Methods and Approaches to Appraisal 60

4.4 Dimensions and Criteria for Appraisal 60

4.5 Sources of Performance Data for Appraisals 61

4.6 Standards for Evaluation of Performance 62

4.7 Special Review and Renewed Probation 64

4.8 Program Evaluation and Maintenance 64

4.9 Summary Conclusions 65

Appendices 67

References 70

TABLES AND FIGURES

Page No.

Part 1

Exhibit 1: General Summary of Performance Appraisal Literature 3

Part 2

Exhibit 2: Choice of Appraiser: Advantages and Disadvantages 15

Exhibit 3: Multifaceted Appraisal System: A Hypothetical Model 25

Part 3

Table 1: Initial Design and Implementation of the Appraisal System 35

Table 2: Current Objectives and Purposes of Appraisal 37

Table 3: Specific Purposes of Appraisals 38

Table 4: Appraisal Design and Techniques 39

Table 5: Standards and Criteria of Performance 41

Table 6: Sources of Performance Appraisal Information 43

Table 7: Training and Preparation for Appraisals 46

Figure 1: Example of a Performance Monitoring System with Special Performance Review 50

ACKNOWLEDGEMENTS

The authors of this study wish to acknowledge the significant contributions of all field site personnel whom we interviewed for our study. Without their assistance and co-operation this project would not have been possible. Special recognition should be given to Dr. Glen Tecker whom we interviewed at Educational Testing Services, Princeton, New Jersey. His insights into the process and methods of appraisal were both inspiring and highly informative.

The authors would also like to acknowledge the guidance and assistance of the Advisory Committee of the Ministry of Education, under the chairmanship of Dr. Peter Nightingale, over the two years of work involved in this project. The comments and assistance received from the members of the Advisory Committee were a crucial asset to the progress and completion of the project.

ABSTRACT

This research project consisted of two major parts: a) a review of the literature on performance appraisal with special attention to the literature on certificated personnel in public education systems; b) field site studies of performance appraisal in action at ten professional service organizations in Ontario and seven public school systems in the USA. Each of the sites studied was selected on the basis of reputation for advanced practices in appraisals.

The literature review provided two major approaches to conducting performance appraisals: a) a focus on the "process" of professional work or the techniques, means, skills, and styles used to perform professional service; and b) a focus on the "outputs" of professional work, or various measures, outcomes, and judgements on the results of professional work. Secondly, the literature review provided two major uses of appraisals: a) "formative" uses where appraisal information is used for recognition, guidance, coaching, development, or problem correction in professional work; and b) "summative" uses, where appraisal information is used for administrative decisions on personnel such as compensation, promotion, training, transfers, or discipline. These approaches and uses of performance appraisal are not mutually exclusive and various combinations can be found in many systems.

Finally, another major theme in the literature review was the growing emphasis on protection of employee rights legislation and the vulnerability of employers to lawsuits arising from contentious personnel decisions based on performance appraisals. These trends have led to considerable effort to improve the reliability and validity of appraisal systems, but have also led to increased bureaucratization of procedures.

The field site studies mirrored many of the results from the literature review, with three important additions:

a) many appraisal systems are in constant evolution so that what is reported in the literature is sometimes not in practice when the site is visited;

b) different approaches and priorities in appraisal are used depending upon the local environmental situation of the organization;

c) bureaucratic cumbersomeness of many appraisal systems causes short-cuts to be used. In practice, this means that most teachers/professionals receive at least "satisfactory" ratings and the elaborate quasi-judicial safeguards built into many systems are used only rarely.

While no generally suitable appraisal system was found to emerge from either the literature review or field site study, two principles seemed to be operative in the most satisfactory sites:

a) under a set of general principles or guidelines established by a relevant central authority, a decentralized appraisal system is developed in operational detail by the local organization to reflect local concerns and priorities;

b) the decentralized system is developed jointly by supervisors and subordinates in the organization so that all involved parties accept the relevance and legitimacy of the appraisal system. Locally developed systems, which satisfy overall standards set by the larger organization, seem to be best suited to working through the inevitable learning curve of implementation and adaptation of an effective appraisal system.

1. INTRODUCTION

This report summarizes the findings of a contract research project authorized by the Ministry of Education of Ontario as one of a series of research studies on performance appraisal in the public education system. The purpose of this study was to conduct a review and analysis of published literature on performance appraisal in general and particularly on the use of appraisals in public education systems. Secondly, the study consisted of a series of field site investigations of performance appraisal systems in action to compare actual practice with results reported in the literature. Two types of field sites were chosen for the study: a sample of organizations in Ontario which employed a large proportion of professional service employees such as researchers, consultants, and educators, and a sample of public school boards in the United States which had a reputation for advanced practices in appraisal of teachers. The results from the field studies provided an opportunity to compare the information from the literature with appraisal practices in organizations which might be analogous to public school boards in Ontario. In this way, the vast range of information on performance appraisal could be narrowed or focused on information relevant to the needs of the public education system in Ontario.

In summary, the specific objectives of the study were:

i) to review the state-of-the-art theory and practice of performance appraisal approaches and methodologies;

ii) to identify those purposes of performance appraisal that might be appropriate to an educational system; and

iii) to describe those approaches which may be particularly relevant to the needs and practices of Ontario's educational system.

The results from this study are intended to provide a common base of information for use by the parties in Ontario most responsible for, and most affected by, performance appraisal of certificated educational staff. To insure that all parties shared the same perception of the nature of the study and the same access to results of the study, an Advisory Committee was established to oversee the definition and progress of the study. The Advisory Committee was composed of representatives from the Ministry, trustees, and professional associations involved in public education in Ontario, as well as representatives from other related contract research projects. Periodic progress reports by all the contractors were presented and discussed in Advisory Committee meetings to ensure co-ordination between projects and to obtain advice and commentary from the committee.

Research into the appraisal and evaluation of educational personnel reveals a broad-ranging and often complex set of issues that influence the acceptability and effectiveness of any system. Part two of this report presents a summary of the literature intended to highlight the individual and organization-level choices which must be exercised to implement or revise appraisal in any school system.

While the literature is replete with descriptions of the controversy, inconclusiveness, and complexity of teacher appraisal and evaluation, it is these features of performance appraisal which have helped to stimulate a body of research and commentary that offers the educator both a range of options in designing appraisal and greater sensitivity to the differing needs of participants to the process. The literature summary is organized into five categories of teacher performance appraisal: (1) Purpose; (2) Components; (3) Process; (4) Outcomes; and (5) Link to Administrator Evaluation. These categories are synthesized around the key questions that emerge in designing or revising an appraisal system. Exhibit 1 shows the five major topics and related sub-topics covered in part two of the report.

Exhibit 1

General Summary of Performance Appraisal Literature

DIMENSION	APPRAISAL ISSUE	KEY QUESTION(S)	DECISION OPTIONS
PURPOSE	Philosophy	. Why do we need to evaluate?	A. Accountability B. Improvement C. Supportiveness
	Objectives	. What do we specifically want to achieve with evaluation?	A. Formative Evaluation B. Summative Evaluation
COMPONENTS	Type of Criteria	. What makes a teacher effective?	A. Presage Criteria B. Process-Oriented Criteria C. Output-Oriented Criteria
		. Why is it so hard to define teacher effectiveness?	A. Complexity B. Context
	Method of Measurement	. How can we actually tell if the teacher is effective?	A. Personality Traits B. Performance/Behavior C. Skills/Competencies D. Teacher Tests E. Goal-Setting F. Clinical Supervision G. Goals-Free Evaluation
	Data for Measurement	. Where should the information for assessing effectiveness come from?	A. Classroom Observation B. Non-Classroom Activities
PROCESS	Choice of Appraiser	. Who should evaluate performance?	A. Supervisor B. Peers C. Self D. Students E. Expert Outsiders
		. How many people should evaluate?	A. Single Appraiser B. Multiple Appraisers C. Interactive Appraisal
	Appraisal Instruments	. What kind of documents should be used?	A. Checklists B. Narratives C. Rating Scales D. Goal-oriented Instruments
	Appraisal Interviews	. How should information be fed back to the teacher?	A. Process of Interviewing
	Appraisal Training	. Should training in the conduct of appraisal be provided?	A. Inservice B. Education Consultants
	Assessment Period and Timing	. How often should appraisal be conducted?	A. Observations B. Conferences C. Assessments
		. Should timing be fixed or based on need?	A. Scheduled Appraisals B. Appraisal in Response to Problems
		. Is evaluation a continuous or discontinuous process?	A. Day-to-Day Feedback B. Periodic Data Collection
OUTCOMES	Evaluation of Outcomes	. How do we evaluate the evaluation system itself?	A. Systematic B. Impressionistic
	Congruence of Outcomes with Stated Purposes	. Is appraisal doing what we wanted it to?	A. Acceptability B. Validity C. Satisfaction D. Accountability
LINK TO ADMINISTRATOR EVALUATION	Interactive Appraisal Design	. Should teacher appraisal be linked to administrator evaluation?	A. Separate Appraisal Systems B. System of Appraisal

Part three of this report presents results from field site studies of selected appraisal systems in actual practice. The field sites were grouped into three subgroups as follows:

i) a Business group composed of five profit-based or self-financing organizations providing professional services such as research, consulting, and teaching;

ii) a College group composed of five colleges of applied arts and technology (CAAT) in Ontario which are post secondary, tuition-based educational organizations; and

iii) a US Schools group composed of seven leading public school systems and one training and consulting organization for teacher evaluation.

All of the above organizations are analogous to Ontario public school systems in various ways and of course are different in various ways. These aspects of the field site sample will be discussed in greater detail in the section of the report which concentrates on the field site part of the research project. In general, the study of analogous organizations focused upon the following issues:

i) to identify the appraisal methodologies and the uses of appraisals in the various types of organizations;

ii) to identify major differences (if any) in approaches to appraisal which seem to be required by the different organizational types;

iii) to identify differences in approaches to appraisals in different functions of the organization and in different levels in the hierarchy;

iv) to identify whether process or output approaches to appraisals are typically used in some functions and organizations and not in others;

v) to identify various approaches to developing criteria and gathering data on the work of professionals; and

vi) to identify linkages between the appraisal and evaluation process and other parts of human resource management in the organization such as compensation systems, training and development, career planning, and supervision.

The final part of the report provides an integration of the results from the various investigations in the study as well as the authors' conclusions on the results. Since this study did not look directly at public education systems in Ontario (this was done in other related research projects), specific recommendations and conclusions for Ontario are not provided. Instead the authors have followed the requirement to provide a common information base to participants in the development of policy for the educational system in Ontario

and it will be these persons who eventually develop specific conclusions and recommendations for use in Ontario. This report provides, to the best of our judgment, a set of generally applicable conclusions for an effective system of performance appraisal in education and it is hoped that these will provide a useful foundation for guidelines for future development.

2. LITERATURE REVIEW: A SUMMARY

2.1 PURPOSES OF APPRAISAL

The purposes of appraisal incorporate both the philosophy of the appraisal system and specific objectives. The philosophy of appraisal pertains to the question: why do we need to appraise and evaluate? Differing perspectives on the philosophy of teacher appraisal are implicit rather than explicit within the literature and uncover important beliefs, values, and assumptions surrounding the issue of performance appraisal. Among the most important of these are varying beliefs concerning:

(1) the significance of teacher effectiveness for the promotion of learning in relation to other variables (student effectiveness, principal effectiveness, school climate, policy, material support, for example);
(2) the potential for individual improvement within a supportive context;
(3) the degree of autonomy and self-determination necessary for effective teaching;
(4) the role of the principal in educational leadership and supervision;
(5) the potential for trust and interactive effort among different levels in the educational hierarchy; and
(6) the extent of co-ordination between instruction and administration in educational organizations that is necessary for student enhancement.

Although the literature does not explicitly address the notion of a philosophy of appraisal, it is revealed by recurring statements about the overall purposes of appraisal. These include:

(1) the promotion of personal and professional growth of employees;
(2) provision of a valid basis for making administrative decisions;
(3) improvement in instructional or school effectiveness;
(4) promotion of student progress and growth;
(5) improvement in the general quality of education; and
(6) accountability to stakeholders in the educational process (teachers, administrators, students, parents, the general public, and so on).

The importance of establishing a clear and acceptable philosophy of appraisal at the outset of system design is supported by two phenomena which pervade discussions in the literature:

(1) the agreement among most researchers and practitioners that some form of assessment is indispensable to the improvement of the quality of education; and
(2) the anxiety, discomfort, and, in some cases, resistance that currently attends the evaluation process for many teachers and administrators.

While there is fairly strong support in the literature for using evaluation, in general, to improve instruction, the specific objectives of appraisal are more susceptible to controversy.

The objectives of appraisal pertain to the question: what do we specifically want to achieve with appraisal and evaluation? Here, the literature tends to make strong distinctions between formative appraisals and summative evaluation. "Formative" appraisal refers to the use of appraisal for purposes of teacher development, feedback, and the identification of strengths and weaknesses. "Summative" evaluation, by contrast, refers to the evaluation of performance of individuals for purposes of making administrative or personnel decisions (e.g., employment, tenure, promotion, dismissal). Other sources of divergence identified in the literature, which address the difficulties of designing an acceptable appraisal system, are related to this distinction between formative and summative evaluation. These include:

(1) the paradoxical role of supervisor as judge versus helper;
(2) the issue of "accountability" versus "improvement";
(3) the "developmental" versus "evaluative" aspects of appraisal;
(4) the issue of teacher autonomy versus control and supervision; and
(5) the need to gather (and feed back) information on specific individuals versus the need to gather a broad sample of information for defining common standards among individuals.

Most of the educational literature speaks to the formative aspect of appraisals. The management and personnel literatures, by comparison, are more oriented to the use of appraisal for evaluation as well as motivational purposes.

The tendency for educational researchers and practitioners to de-emphasize summative evaluation has its roots in several factors:

(1) teacher behaviour is relatively autonomous and difficult to observe;
(2) the "outputs" or results of teaching are hard to assess;
(3) the teaching process requires considerable discretion and self-directed creativity in order to be effective;
(4) unlike summative evaluation, formative appraisal does not require the identification of common generalizable criteria for purposes of comparison among individuals;
(5) the organizational context of instruction (e.g., physical separation, high teacher-principal ratio) impedes informal performance feedback;
(6) personnel decisions in education organizations are often linked to seniority considerations;
(7) summative evaluation creates more anxiety among appraisers and appraisees than formative appraisals; and
(8) formative appraisal avoids the task of balancing the stated needs of various stakeholders in the education process, since formative appraisal is primarily teacher-focused.

The conceptual and practical division between formative appraisal and summative evaluation reflects a traditional division within the literature (and in practice) between the purpose of appraisal for improvement of performance versus judgment of performance. Recent literature, however, points to trends within the United States social and political system that underscore a growing emphasis on summative evaluation. For example:

(1) strong public response to the findings of the 1983 report of the National Commission on Excellence in Education;[1]
(2) growth in state and local school system adoption of teacher competency testing as a condition of certification;[2]
(3) increases in the pace of state-mandated teacher evaluation (in 1983, 26 states required teacher evaluation as compared to only six states in 1970); and
(4) public resurgence of the merit pay controversy.[3]

Although none of the current literature offers specific approaches for reconciling the formative-summative dichotomy, researchers and practitioners provide some general suggestions for combining the purposes of appraisal:

(1) teachers should be major participants in the design and process of appraisal;
(2) methods of appraisal should be part of in-service training for both administrators and teachers;
(3) day-to-day feedback and support needs to be an integral part of the overall appraisal system;
(4) one of the criteria in administrator assessment should be his or her effectiveness in appraising and evaluating teachers;
(5) the context of teaching should be understood so that teachers are only evaluated on the basis of behaviours under their control;
(6) the purposes and processes of appraisal and evaluation should be widely communicated (e.g., manuals, committees) and easily understood;
(7) some system of recourse or appeal must exist for complaints and grievances pertaining to the content and process of evaluation; and
(8) systems should be designed which are interactive, democratic, and mutually supportive rather than autocratic and control-oriented.

Overall, the literature assumes that resistance to and inconsistencies between formative appraisal and summative evaluation are minimal when the appraisal system is simple, valid, and equitable.

---

1 See, for example: "Can the Schools Be Saved?" Newsweek, May 9, 1983.
2 Hathaway, W.E. "Testing Teachers"; Educational Leadership Vol 33 (1980), pp. 210-215.
3 See, for example, Toll, S. "This merit pay plan for management improves performance and sets an example for other employees (like teachers, for instance)". American School Board Journal Vol 170, no. 8 (August 1983), pp. 38-9; or Rist, M.C. "Our nationwide poll; most teachers endorse the merit pay concept". American School Board Journal Vol 170, no. 9 (September 1983), pp. 26-33.

2.2 COMPONENTS OF APPRAISAL

The components of appraisal can be summarized around three major topics:

(1) type of criteria;
(2) method of measurement; and
(3) data for measurement

2.2.1 Type of Criteria

Choice among types of criteria involves the question: what makes a teacher effective? Lack of conclusive evidence on teacher effectiveness is attributed to a host of factors in the research literature. In general, the difficulty of determining what makes a teacher effective pertains to two issues, complexity and context, both of which inhibit the identification of commonalities among teachers for the prediction of effectiveness. Complexity refers to the availability and exercise of diverse means (teacher behaviours, activities, and approaches) for the attainment of a single end (pupil progress). Most arguments in the literature surround two distinct aspects of complexity: the degree of complexity, diversity, or ambiguity inherent in the teaching process,[4] and the degree of discretion and flexibility there should be. While there is a general consensus within the literature that the teaching process is complex, the greater the number of appropriate teaching methods (that either do exist or should be permitted) for achieving the single end of student growth, the greater the difficulties in establishing common criteria that meaningfully distinguish between good and marginal teachers for purposes of appraisal. The second problem in identifying valid criteria of teacher effectiveness is context. The causal link between teacher effectiveness and student outcomes is moderated significantly by a large number of contextual factors; what is more, the presence and influence of these factors are likely to vary considerably among teachers. Context factors include:

(1) student characteristics (e.g., grade level, motivation, socio-economic status);
(2) instructional setting (e.g., materials and professional resources, structure of curriculum);
(3) school climate, policy and organization;[5] and
(4) the social context and conflicting expectations of different educational constituencies.

---

4 See, for example, Darling-Hammond, L.; Wise, A.E.; and Pease, S.R. "Teacher Evaluation in the Organizational Context: A Review of the Literature". Review of Educational Research Vol 53, no. 3 (Fall 1983), pp. 285-328.

5 See, for example, Lotto, L.S. "Revisiting the Role of Organizational Effectiveness in Education Evaluation". Educational Evaluation and Policy Analysis Vol. 5 no. 3 (Fall 1983), pp. 367-378; or Weick, K.E. "Administered Education in Loosely Coupled Schools". Phi Delta Kappan Vol. 63, no. 10 (1982), pp. 673-76.

All of these variables tend to limit the number of criteria of teacher effectiveness which are generalizable across educational settings.

There are three types of criteria for assessing teacher effectiveness:

(1) presage criteria: the characteristics with which the teacher comes to the job, including personal and professional characteristics;

(2) process-oriented criteria: criteria based on the process of teaching, including teaching methods, behaviours, and techniques, student-teacher relationships, and classroom strategies; and

(3) output-oriented criteria: criteria based on student learnings, pupil outputs, and changes in pupil behaviour.

A large number of advantages and disadvantages adhere to each of these types of criteria depending upon the objectives and content of the appraisal program. Nevertheless, some tentative conclusions from the literature appear warranted at this time:

(1) presage criteria are more applicable for use in the recruitment and selection of teachers than as standards to be utilized in ongoing assessment;
(2) presage criteria which correlate with teaching effectiveness are: scholarship in college, professional preparation, attitude toward teaching, age, and experience;
(3) presage criteria which appear uncorrelated with teaching effectiveness are: knowledge of subject matter, cultural background, intelligence, socio-economic status, sex, and marital status;[6]
(4) process-oriented criteria are more common and easier to use than output-oriented criteria; the use of output-oriented criteria tends to create resistance to evaluation; and
(5) the validity of output-oriented criteria suffers from the inability to control for all context variables that moderate the teacher behaviour-pupil learning relationship.

2.2.2 Method of Measurement

Whereas choice of criteria attempts to answer the question "what makes a teacher effective?" choice of method of measurement asks: how can we actually tell if the teacher is effective? Seven methods emerge from a review of the literature; these are summarized below.

(1) Traits (Personality Characteristics)

---

6 See Farquhar, R.H. "How the Teaching Profession Measures Teaching Effectiveness". 1978 (ED 1985 683); or Feldvebel, A.M. "Teacher Evaluation: Ingredients of a Credible Model". Clearing House, May 1980, pp. 415-420.

Assessment of teachers on the basis of personality traits tends to produce measures which are ambiguous or non-specific (e.g., "the teacher is conscientious", or "enthusiastic"); this ambiguity makes the process of appraisal much more prone to subjectivity and appraiser bias. Also, because this type of measurement lacks behavioural specifics, teachers find it hard to determine how to improve their performance. Despite their continued use, assessment on the basis of traits is usually unreliable and inaccurate.

(2) Performance/Behaviour

This method of measurement examines the actions or behaviours which are expected to occur if performance is satisfactory. Performance based or behavioural data possess the greatest potential for validity, reliability, discriminability, and receptivity. The measurement of specific behaviours also allows for concrete and meaningful feedback to the teacher. The development of behaviour-based standards is more time-consuming than the trait approach, but existing instruments[7] can be borrowed and revised to accommodate individual district or school needs.

(3) Skills and Competencies

Unlike behaviour-based measurement of performance, assessment of competencies or skills only measures capacity. That is, skills-based assessment measures what the teacher is able to do rather than the extent to which this capacity is applied consistently or appropriately. This distinction is commonly overlooked in the education literature. Considerable progress has been made in developing a valid inventory of competency statements that describe effective teaching.[8]

(4) Teacher Tests

Although 15 U.S. states now require some form of tests for teachers as a condition of certification (e.g., The National Teacher Examinations), the use of teacher tests has not been applied to the appraisal of certificated personnel and is unlikely to be used as a basis of measurement in the future. Despite a Supreme Court ruling that NTE scores are related to teacher job performance, the validity of teacher testing is still a source of controversy in the literature.

(5) Goal-setting (MBO)

Unlike measurement against a predetermined set of criteria, goal-setting involves the assessment of performance against mutually derived goals and objectives. Research and commentary on the advantages and disadvantages of goal-based appraisal are numerous. For

---

7 For suggested sources, see Lewis, A.C. Evaluating Educational Personnel. American Association of School Administrators, 1982.

8 See for example, Gutridge B.M. Teacher Competency: Problems and Solutions. American Association of School Administrators, 1980.

example, goal-setting tends to be time-consuming, provides limited feedback on specific teacher behaviours, and requires consistent commitment by all participants in order to be useful. On the other hand, goal-setting is an interactive process which encourages hierarchical communication, self-direction, self-evaluation, and co-operation between teacher and supervisor. There is no conclusive evidence at present on the acceptability and validity of goal-setting.[9] Several sources advocate its use in combination with predetermined criteria.

(6) Clinical Supervision

Clinical supervision covers a fairly broad literature that describes a democratic, interactive approach to the improvement of instruction. Clinical supervision consists of three major activities: planning conferences, classroom observation, and feedback conferences. The objectives of clinical supervision are professional development, objective feedback, skill development, and the diagnosis of instructional problems. Its use as a basis for making personnel decisions (summative evaluation) has been suggested but not explored in any detail in the literature. The links between clinical supervision and teacher performance, and between clinical supervision and student performance, have not yet been convincingly demonstrated.[10]

(7) Goal-Free Evaluation

A final method of measurement which departs significantly from the foregoing approaches is "goal-free evaluation".[11] This approach argues against the use of any specific, standard set of criteria. Instead, teachers "make their best case for merit or value" based on multiple lines of evidence that reflect the needs and understandings of all audiences who use the evaluation results. These "lines of evidence" may in fact constitute, among other things, several methods of measurement or types of criteria described earlier (e.g., performance tests, personal characteristics, credentials, competency-based evaluation). Deficient teachers are identified as those who are unable to provide credible or sufficient evidence to be considered to have an adequate case. The validity of this method of measurement has yet to be assessed. Although goal-free evaluation overcomes the problem of identifying commonalities among teachers for purposes of comparison and attempts to respond to the needs of all stakeholders in the process, this method is more vulnerable to evaluator bias and subjectivity in making summative assessments.

---

9 For a thorough treatment of this approach, see Redfern, G.B. Evaluating Teachers and Administrators: A Performance Objectives Approach. Westview Press, Boulder, Colorado: 1980.

10 Acheson, K.A., and M.D. Gall, Techniques in the Clinical Supervision of Teachers. New York: Longman, 1980.

11 Peterson, K., and D. Kauchak, "Progress on Development of Lines of Evidence for the Evaluation of Public School Teaching". Paper presented at the American Educational Research Association, Montreal, Apr. 1, 1983 (ED 228 228).

2.2.3 Sources of Data For Measurement

The third component of appraisal concerns data for measurement. This aspect of appraisal is a response to the question: where should the information for assessing teacher effectiveness come from? There are two sources of appraisal data: classroom observation and non-classroom activities. The current literature focuses almost exclusively on the former data source and relates specifically to the "how-to" of observation.[12] Conceptualizations of alternative data sources can be re-cast in terms of who should do the appraising, since choice of appraiser (e.g., outsider, principal, teacher) is the primary determinant of what data are included in appraisal. This brings us to the first of several process factors in the appraisal of teacher effectiveness.

2.3 PROCESS OF APPRAISAL

The process of appraisal refers to the mechanics of how appraisal is conducted. There are five major aspects to the process of appraisal:

(1) choice of appraiser,
(2) appraisal instruments,
(3) appraisal interviews,
(4) appraisal training, and
(5) assessment period and timing.

2.3.1 Choice of Appraiser

Choice of appraiser pertains to two specific questions: who should appraise and evaluate performance? and how many people? A variety of sources in the literature address these issues either directly or indirectly.

(1) Who should appraise and evaluate performance?

Answers to this question are likely to evolve, in practice, from the implicit philosophy of the appraisal system itself. For example, an emphasis on development, growth, and autonomy requires self-evaluation as an integral component of the appraisal process. Choice of appraiser also emerges from the prescribed objectives of appraisal. Where the purpose of performance appraisal is summative, the supervisor is assumed to have responsibility for making the evaluations of the relative performance of teachers upon which personnel decisions will be based. Furthermore, the decision as to who should participate in appraisal is also influenced by the strength of interest and degree of perceived co-operation among different constituents in the education process (for example, the amount and acceptability of input from public officials, parents, and outside experts). Direct participants in the appraisal process

---

12 For a particularly thorough treatment of this topic, see McDonald, F. "Principles and Procedures in Observing Classroom Instruction". In Peterson, D., and A. Ward, eds., Due Process in Teacher Evaluation. Washington, D.C.: University Press of America, 1980, pp. 110-113.

may include: (1) supervisor; (2) self; (3) peers; (4) students/subordinates; or (5) expert outsiders. There is fairly extensive research and practical commentary on the efficacy of these alternative sources; a summary of advantages and disadvantages is outlined below in Exhibit 2.

All appraisal sources are prone to some degree of bias and no research to date has provided conclusive evidence on the relative empirical validity of these five sources. Nonetheless, investigations into methods for reducing rater error have been most vigorous in the area of supervisor appraisals, since this form of assessment continues to be the most common.

(b) How many people?

The question of how many people should appraise and evaluate is important for several reasons. First, the decision to use only one appraiser precludes the opportunity for making appraisal an interactive, mutually supportive process. In this way, the choice of appraiser becomes a manifestation of the philosophy which underlies the system. Second, research indicates that the use of multiple sources of evidence tends to increase the validity and reliability of the appraisal results and reduce judgment error or bias. For this reason, some schools experiment with a "team approach" to appraisal. The benefits of increased validity, however, have to be considered in light of two potential disadvantages: the use of multiple appraisers is more costly and time-consuming, and the process may be threatening to the appraisee. Finally, knowing to whom the appraisal data will be made available also affects the number of data sources chosen. For example, teachers are less resistant to formal input from students, outsiders, and peers if such information is fed back only to the teacher. The proposed audience for appraisal data is, in fact, a reflection of the purposes that the appraisal system is intended to serve. Formative appraisal does not require teachers to share data with others (although it is, naturally, encouraged). The validity of summative evaluation, by comparison, increases in proportion to the amount of data or number of sources available for making judgments about teacher effectiveness.

2.3.2 Appraisal Instruments

Appraisal instruments refer to the specific documents used in evaluating teacher effectiveness. There are only limited examples of actual appraisal instruments in the educational literature [13] but common instrument characteristics tend to vary among four overa designs. These are summarized below.

---

13 See, for example, Gutridge, B.M. Teacher Competency: Problems and Solutions, op. cit.; Lewis, A.C. Evaluating Educational Personnel, op. cit.

Exhibit 2

CHOICE OF APPRAISER: ADVANTAGES AND DISADVANTAGES

Choice of Appraiser	Advantages	Disadvantages
SUPERVISOR	- supervisor hierarchically accountable for the performance of subordinates - supervisor responsible for personnel decisions - provides a forum for supervisor-appraisee interaction and communication	- may promote discomfort for appraiser or appraisee - prone to bias and subjectivity
SELF	- suitable for developmental purposes - encourages self-growth and motivation - overcomes difficulty of ongoing observation by externals	- unsuitable for purposes of making personnel decisions - tendency toward bias and distortion - low predictive validity for student achievement
PEER	- fellow professionals provide relevant valuable input - peers more apt to take into account day-to-day job demands	- peer group pressures may create bias or distortion - may create disharmony or competition - timetable restrictions impede its practicability
STUDENT/ SUBORDINATE	- suitable for developmental purposes - provides valuable feedback to appraisee	- inconclusive evidence on predictive validity - age of student restricts its use
EXPERT OUTSIDER	- increases validity and objectivity of appraisal - provides valuable 'expert' feedback - useful for resolution of appraiser-appraisee conflict - saves supervisory time	- may be expensive - process may be threatening to appraisee - use of third party discourages meaningful interaction about performance improvement between supervisor and subordinate

(1) Checklists

Checklists are comprised of a list of statements (describing specific traits or behaviours) beside which the appraiser is asked to check which are applicable to a particular employee or observed by the appraiser. Checklists indicate the presence or absence of a particular attribute without reference to the degree to which that attribute prevails. During observation, a checklist is simple to use and directs attention to aspects of a lesson which an appraiser might otherwise miss. On the other hand, classroom observation may become a mechanical, routine procedure in which supervisors are apt to make judgements without patient reflection or careful analysis.

(2) Narratives

The narrative technique (or "essay evaluation") is a written report of the teacher's strengths, weaknesses, or future potential. The rich descriptive nature of this method makes it a worthy supplement to other types of techniques for purposes of feedback. Narratives are of limited use in isolation, however, as a basis for making judgemental decisions that involve meaningful comparison among teachers.

(3) Rating Scales

The most widely used evaluation technique involves judgemental rating of various characteristics of the teacher's performance along some graphic scale or continuum. The scale can range in specificity from a simple "unsatisfactory - needs improvement - satisfactory" to more specific descriptions of different levels of performance or exhibited behaviour (e.g., behaviourally anchored rating scales). Behavioral rating scales are more useful for appraisal-feedback purposes than simple judgemental rating scales which tend to be used more for evaluation purposes.

(4) Goal-Oriented Instruments

This type of instrument is used when the method of measurement for appraisal and evaluation is goal-setting. Although the format of the instrument depends on the individual designers or users, goal-oriented instruments usually provide space for one or more of the following: (1) identifying the goals to be achieved; (2) specifying how the goals will be accomplished; (3) describing or rating the extent to which goals have been achieved; (4) providing evidence of progress toward goals; and/or (5) identifying personal or organizational factors which impede or facilitate goal accomplishment. Goal-oriented instruments provide a rich amount of appraisal data which can also be used for evaluation purposes.

The choice of which technique to use is contingent upon its purpose (formative appraisal versus summative evaluation) and the appraisal designer's individual assessment of which benefits and drawbacks inherent in each technique are most relevant to the particular context.

An alternative approach is instrument design that combines more than one of the foregoing techniques. Because hybrid techniques overcome several of the weaknesses of any single approach used in isolation, instruments are sometimes designed, in practice, to include both goal statements and performance-based rating scales.

2.3.3 Appraisal Interviews

The choice of appraiser and type of instrument are concerned with where data should come from and how information should be collected. Appraisal effectiveness also depends on how the information is fed back to the teacher; a good appraisal interview (or conference) is both systematic and humane. To a large extent, the discomfort which often attends the appraisal interview is a function of the quality of (and participation in) decisions made on all other performance appraisal issues. For example, the appraisal interview is likely to be less stressful if:

(1) both the supervisor and teacher have a clear and compatible understanding of why the evaluation has been conducted (purpose of appraisal);
(2) the teacher knows what is expected of him or her (criteria and standards);
(3) the teacher expects that the process will be supportive (articulation of appraisal philosophy);
(4) the teacher knows in advance how the appraiser has assessed performance (data for measurement and appraisal instruments); and
(5) the teacher has full knowledge of where the information has come from (appraisal sources) and how it will be used (objectives of appraisal).

The educational literature suggests that while appraisal conferences are acknowledged as a key contributing factor to successful appraisal, little research has been conducted to guide practitioners through the process.[14] The management and personnel literatures, by comparison, have generated a broad range of advice and instruction on how to conduct an appraisal interview. Appendix 1 provides a synthesis of the bulk of the more recent literature on this topic. Overall, the keys to successful interviewing are careful preparation, open dialogue, and a problem-solving rather than fault-finding climate.

2.3.4 Appraisal Training

Inservice training in the process of performance appraisal (for both appraisers and appraisees) may be another major determinant of successful appraisal. To date, both research and practice have tended to neglect this issue. Yet there are several potential benefits of inservice training on the "how-to" of appraisals. For example, formal training programs may:

(1) increase both the appraiser's and teacher's confidence or faith in the appraisal process;
(2) clarify the purposes of appraisal;

---

14 Sweeney, J. "The Post-Observation Conference: Key to Teacher Improvement". The High School Journal Vol 66, no. 2 (1983), pp. 135-140.

(3) provide a vehicle for disseminating information on appropriate appraisal methods;
(4) reveal unintentional rater biases;
(5) promote consistency and a common frame of reference for the conduct of appraisal; and
(6) encourage involvement and participation among hierarchical levels within the school organization.

While it is too early to assess the effectiveness of such endeavors, a small minority of schools and school districts in the United States are beginning to develop successful inservice training programs or hiring experts in the field of education to conduct workshops and seminars on the appraisal process.

2.3.5 Assessment Period and Timing

The fifth and final dimension of the appraisal process pertains to three distinct questions related to assessment period and timing. First, how often should assessments be conducted? There is currently no specific research on the optimal number of times per year that assessment should occur. The more often that data are collected, the more timely the information and, therefore, the more accurate. However, time, cost, and convenience considerations usually restrict the frequency of both observations and evaluations, in practice, to once a year or once every few years.

The second question which deserves attention is: should timing be fixed or based on need? Again, there is no explicit treatment of this question in the literature, but the answer, in practice, commonly reveals the stimulus which drives the evaluation process. That is, some appraisers adhere to a systematic schedule of observation, conferences, and summative assessments, whereas others only evaluate when a critical problem surfaces. In the latter case, supervision becomes "management by exception" and the timing of assessment is driven by the emergence of problems or negative information rather than the positive desire to assist all teachers. Under such circumstances, the supervisor may "waive" observations or formative appraisals for those teachers whom he or she perceives as performing adequately, which in turn denies good teachers the opportunity for positive feedback and sustains an image of the appraisal process as negative or punitive.

There is one final question which relates to assessment period and timing: is appraisal continuous or discontinuous? Whatever interval for assessment is chosen, appraisal can be defined as a continuous process, because information about employee performance is continuously monitored and feedback to the teacher should occur often on a more informal basis. Where the appraisal process is effective, the day-to-day support of the teacher should be a natural outgrowth of the formative appraisal system. Evaluations, however, may be a more periodic event as required by the summative evaluation needs of the system or by particular problems of performance.

2.4 OUTCOMES OF APPRAISAL

If an appraisal system is effective, the outcomes of appraisal are virtually identical to the purposes of appraisal (e.g., student growth, instructional effectiveness, professional improvement, valid personnel decisions, teacher feedback, and support). None of the current literature confronts the evaluation of system effectiveness specifically, that is, how do we evaluate the evaluation system itself? Impressions of effectiveness of appraisal systems tend to be inferred informally from one or more of several indicators; for example: satisfaction, school morale, public response, number of grievances, level of absenteeism, pupil test scores, informal feedback, amount of paperwork, or costliness of the system. One notable exception to the absence of systematic methods for evaluating the appraisal program itself is the Teacher Evaluation Needs Identification Survey.[15] Developed in Hartford, Connecticut, to determine the degree to which a school's evaluation system complies with state guidelines for effective teacher evaluation programs, TENIS comprises a staff survey that assesses the status of teacher evaluation practices in local schools. Results obtained from field tests substantiate the instrument's validity and reliability as well as its sensitivity to differences in teacher evaluation practices.

In general, the outcomes of appraisal concern the question: is appraisal doing what we wanted it to? In the educational literature, outcome issues are addressed within a more conceptual framework that relates to one or more of the following criteria of appraisal effectiveness:

(1) acceptability: the extent to which appraisal participants and outside constituents exhibit trust in, rather than resistance to, the process;

(2) validity: the extent to which the appraisal system actually measures what it purports to measure with minimal bias or error;

(3) satisfaction: the extent to which the system meets the needs of individuals and promotes a positive, meaningful working climate; and

(4) accountability: the extent to which the appraisal system clarifies individual and organizational roles and ensures prescribed responsibility for improvement in the quality of education.

Although these four outcome issues are usually treated separately in the literature, they are actually interrelated, in that: (a) acceptability is a function of the degree to which individuals are satisfied with the system and believe that it is valid; (b) the validity of the system is a function of the degree to which individuals are willing to accept the way that appraisal is conducted and the purposes it serves; and (c) satisfaction with the system is a

---

15 Iwanicki, E.F. "Development and Validation of the Teacher Evaluation Needs Identification Survey". Educational and Psychological Measurement Vol 42, no. 1 (Spring 1982), pp. 265-74.

function of the degree to which the process appears valid and equitable. Resistance to appraisal is less the result of an unwillingness on the part of individuals to be held accountable for their actions than the consequence of breakdowns in the conditions that promote accountability. Generally speaking, actions for which individuals are accountable need to be (a) clearly specified, (b) causally linked to the desired goals of quality education, (c) under the individual's control, and (d) a viable basis for rewarding effort and ability.

2.5 LINK TO ADMINISTRATOR APPRAISAL AND EVALUATION

In comparison with the volume of research and commentary on teacher performance appraisal, the educational literature on administrator appraisal and evaluation is fairly limited. While informal or ad hoc systems of administrator appraisal have been used in practice for some time, it is only recently that efforts have been initiated to examine this area more systematically.[16] To date, current practice (particularly among state-mandated programs in the U.S.) offers more specific and extensive information on the design and conduct of administrator appraisal than that which appears in the educational literature.

The issue of administrator performance appraisal is linked to teacher appraisal for several reasons:

(1) it is difficult to separate the effects of teacher behaviour and administrator behaviour on pupil progress and improvements in the quality of education;

(2) an effective system of appraisal encourages hierarchical co-ordination and accountability among all levels of the school organization;

(3) one of the criteria of administrator appraisal should be the administrator's willingness and ability to appraise teacher effectiveness; and

(4) administrator appraisal and evaluation promotes a better understanding among administrative personnel of teachers undergoing assessment.

The effective design and implementation of administrator appraisal systems is contingent upon an identification of the same appraisal issues (philosophy, objectives, type of criteria, and so on) as those outlined above for teacher appraisal. In addition, administrator

---

16 See, for example: De Bevoise, W. "Synthesis of Research on the Principal as Instructional Leader" Educational Leadership Vol 41, no. 5 (February 1984), pp. 15-20; Duhamel, R.; Cyze, M.; Lamacraft, G.; and Rutherford, C. "The Evaluation of Principals". Education Canada Vol 21, no. 2 (Summer 1981), pp. 20-27; Herman, J.J. "An Up-Front Approach to Administrative Evaluation". NASSP Bulletin (Nov. 1978), pp. 34-42; Lewis, A.C. Evaluating Educational Personnel op. cit.; McCleary, L. "Evaluation of Principals". Theory Into Practice (Feb. 1979), pp. 45-49; Renihan, F.I. "The Evaluation of Principals". Saskatchewan Education Administrator (Dec. 1980), pp. 22-31; Stow, S.B., and Manatt, R.P. "Administrator Evaluation Tailored to Your District or Independent School". Educational Leadership Vol 39, no. 5 (Feb. 1982), pp. 353-357.

appraisal programs confront many of the same challenges intrinsic to teacher appraisal, particularly:

(1) the inability of supervisors to observe behaviour on a day-to-day basis;

(2) the ambiguity of criteria for use in assessment of performance;

(3) the need to accommodate all personnel who have the right to a voice in determining the goals, process, and content of appraisal;

(4) the need to determine the relative contribution and relevance of alternative data sources (e.g., supervisors, self, subordinates, parents);

(5) the problem of reconciling the developmental and judgmental purposes of appraisal; and

(6) the difficulty of distinguishing the effects of uncontrollable variables in the work context from the effects of individual behaviour in assessing performance competency or goal achievement.

The essential elements in appraisal of principals include curriculum development, supervision of personnel, school management, and community relations.[17] School administrators have responsibilities which involve the co-ordination of expectations from politicians, parents, teachers, students, superintendents, and community groups. Flynn (1978), for example, states that "administrators who hold positions in school systems have regarded themselves as being in the centre of an array of demands".[18] Therefore, while one of the objectives of performance appraisal may be to establish consistent and reliable standards for use across different individuals performing similar tasks, some flexibility must be retained to allow each school administrator to respond to the unique circumstances and requirements of his or her particular environment.

There are currently very few validated instruments available for the appraisal of administrative performance. In a review of formal programs for measuring administrative competencies, Walters (1980) compared the four most well-known systems: Educational Leadership Appraisal (ELA), Georgia Principal Assessment System (GPAS), Individual Learning Materials (ILM), and Special Education Supervisor Training (SEST). Walters points out that "only GPAS attaches high importance to appraisal by persons other than the individuals themselves."[19] The primary use of most instrument systems is self-assessment and personal development.

---

17 Deal, T.E.; Dornbusch, S.M.; and Crawford, R.A. "Villains as Victims: Evaluating Principals". Phi Delta Kappan, Dec. 1979, pp. 237-74.

18 Flynn, J.J. "Evaluation of School Administrators". Comment on Education (October 1978), p. 5.

19 Walters D.L. "The Measurement of Administrative Competencies". Phi Delta Kappan, Feb. 1980, p. 424.

Although teacher appraisal and administrator appraisal are conceptually and practically related, the educational literature treats these two topics separately. Future research and practice will need to address the question: how should teacher appraisal be linked to appraisal of administrators? Separate appraisal systems are designed around the necessary differences in tasks, responsibilities, and criteria for assessment between teachers and administrators. A system of appraisal is focused on the potential congruencies and interrelationships between teachers and administrators with regard to the philosophy, objectives, and process of evaluation. Future investigation into the development of a system of appraisal is likely to encourage examination of some critical questions surrounding the evaluation of educational personnel. Appendix 2 summarizes the major questions which future research will need to address.

As research progresses in the field of administrative assessment we are likely to witness an increase in attention to the relationship between teacher and administrator appraisal and the potential for a synergistic system of appraisal that recognizes the interactive qualities of performance in the educational context. The links between teacher and administrator appraisal are primary indicators that student growth is a shared responsibility.

The state of current literature on teacher performance appraisal is less a cause for concern than for optimism. At present, there is considerably more opinion in the literature than validated research, and what research has been conducted, particularly on teaching effectiveness, is inconclusive. In addition, current controversies surrounding the objectives and process of appraisal remain unreconciled. Yet the recent growth in vigorous attention to the topic of appraisal bears witness to the earnest attempts of researchers and practitioners to illuminate all sides of the evaluation question. It is such efforts that generate options in the design of appraisal and allow the development of a flexible appraisal program which considers the needs of all participants to the process. More importantly, the lack of consensus on particular details of appraisal has not obstructed general agreement that the overall values of honesty and fairness form the substratum upon which effective appraisal is built.

2.6 COMPARISON AND CONCLUSIONS

The foregoing review of the literature is a synthesis of research and commentary from both the general management and educational literature on performance appraisal. Both sets of literature are very similar in identifying several demands or requirements which must be satisfied by an appraisal system if it is to be effectively utilized in an organization. The list of possible uses of an appraisal system is essentially the same for both a business and educational setting.

a) The appraisal system can be used to generate data for personnel administration (compensation, training, counselling, and staff planning) and for personnel management decisions (retention, promotion, transfer, discipline). Decisions for personnel administration require appraisal data which basically indicates whether a person's level of performance is sufficiently satisfactory to retain his or her current job (i.e.,

performance at least satisfies the minimum standards). Decisions for personnel management require data from appraisal which distinguish levels of performance from one person to another, so that equitable and valid standards are used for decisions on retention, promotion, or discipline.

b) The appraisal system can be used for supervision and motivation of personnel (probationary review, communication, negotiation, and clarification of assignments and responsibilities, goal setting, delegation, support, and correction). Supervision of personnel via the appraisal system is accomplished through the communication and discussion of work assignments formally in a post-appraisal discussion and informally throughout the year. If a subordinate knows that these assignments will be subsequently reviewed in the next appraisal session, the motivation to concentrate on these assignments should be enhanced. In addition, if the appraisal system provides an opportunity to participate in the setting of assignments so that the persons are working essentially toward their own goals, then motivation should be enhanced as well.

c) The appraisal system can be used for introduction, maintenance, and/or changes in operating methods (clarification of operating methods, monitoring of methods, identification of training needs to change methods). If the organization is convinced that a particular philosophy or particular methodology of professional service is superior or desirable, then the appraisal system performance criteria can be designed to elicit and reward this particular behaviour. For example, if an educational institution believed that highly theoretical academic research was the path to organization excellence, then the appraisal system would use publications in refereed journals as a performance criterion. If the institution believed that teaching quality was the path to excellence, then student ratings of teachers and courses would be a criterion. If an organization wished to change its philosophy or methodology of professional service, then it would be necessary to develop a new set of performance criteria in the appraisal system to elicit and reward the new type of professional behaviour.

d) The appraisal system can be used for achievement of organizational and personal goals by translation of organization goals and plans into a series of individual goals and plans at all levels of the hierarchy. This approach to appraisal is commonly used where an organization has developed a mission or strategic plan that requires co-ordinated effort toward the achievement of organization goals. Departmental and individual goals can be made up of subsets of the organization goals, or individual task assignments can be derived from the larger goals. In either case, if individuals are held accountable for performance which contributes to the achievement of organization goals, then the likelihood of organization goals being achieved is enhanced.

e) The appraisal system can be used to protect the rights of individuals and provide due process for personnel decisions (fair and open criteria for appraisal, equity of treatment, equal rights or avoidance of discrimination, rights to negotiate and appeal). This use of appraisal is based on assumptions of validity in the appraisal system and open communication of performance standards and criteria. Open communications ensure

that everyone is fully aware of the relevant standards of performance and will not be subject to arbitrary or biased standards. Validity ensures that people are appraised on their actual level of performance rather than on imagined or prejudged levels of performance. Such a system would not allow inadequate performers to hide, but more importantly would give full rights to recognition for competent performers who might otherwise not be recognized. These features of an appraisal system would go a long way towards ensuring equal rights and avoidance of discrimination in an organization.

f) The appraisal system can be used for professional development of groups and individuals in the organization. For individuals, this might come from feedback on their performance and discussions on how to improve (or how to continue) in the future. These discussions might even generate a set of action plans or personal goals on which to be appraised in the next evaluation. Training or development needs for groups may not necessarily be identified from the performance appraisal, but progress of a group toward some development goals could be assessed by aggregating some of the appraisal information. Development of individuals can also emerge from some of the other uses of appraisal, for example, progress toward learning, the organization's preferred style of professional service, or the opportunity to participate in goal setting.

Of course, not every appraisal system will use all these potential applications. This will be dependent on the needs or priorities of a particular organization. Secondly, different purposes or priorities for the appraisal system may exist for different levels in the hierarchy or different types of work within the same organization. Once the priorities and purposes of appraisal for different parts of the organization have been determined, then attention can be drawn to the methods of appraisal. It is clear from the review of both sets of literature that method of appraisal is a complicated and frustrating topic because all methods involve various compromises of cost, convenience, and complexity. The requirements for good appraisal methods are generally well understood, but the requirements are so demanding as to be practically impossible to satisfy by any one method of appraisal. Consequently, we are led to a conclusion that an organization must consider using a multifaceted approach to appraisal methods.

While a multifaceted approach is also complicated, it is possible to reduce some of the complexity by using different methods of appraisal for different types of teachers or positions at different levels in the hierarchy of a school. By way of illustration, Exhibit 3 presents a hypothetical example of how appraisal methods might vary in one school for different types of teachers. Each of the various methods in a multifaceted system must satisfy the criteria of validity, reliability, and acceptability, but these criteria can actually be enhanced by using different appraisal methods for different types of work. Moreover, a multifaceted approach avoids the problem of burdening everyone in the hierarchy with having to go through the administrative bureaucracy of appraisal methods which are not appropriate for different types of work.

Exhibit 3

Multifaceted Appraisal System: A Hypothetical Example

Type of Teacher	Focus of Appraisal	Evaluation Criteria	Possible Methods of Appraisal
1. Entry level or probation	decisions on retention or training	skills, traits knowledge, potential	credentials; rating and behaviour scales; incidents; simulations; possible ranking
2. Experienced permanent staff	maintain standards, professional development	continued teaching effectiveness, non-classroom work, development needs	observations; peers; secondary indicators; activities; negotiated plans
3. Senior teaching staff	maintain standards, wider responsibility	continued teaching effectiveness, administrative or leadership behaviour, results of plans	observations; negotiated work planning and review; results-focused appraisals.
4. Special performance review	employee assistance, upgrading performance standards, dismissal	specific to problems, symptoms or indicators, specific timing and activity plans	negotiated plans, activities and targets; specific observations; trial periods; due process mechanisms.

For example, it may be futile to require a novice teacher to go through an elaborate goal-planning activity for appraisal when the primary concern is whether the teacher possesses minimum acceptable standards of knowledge and skills for teaching in a subject area. Similarly, for an experienced teacher with a long history of excellence in teaching, it may be futile to force the person and the principal to go through a charade of detailed goals-planning when the primary concern is to show recognition and appreciation for good work. On the other hand, for a senior teacher who wishes to qualify for higher levels of responsibility, a detailed career development plan would be quite appropriate both for professional development and for performance appraisal. Thus a contingency approach seems to be the most suitable for the question of choice of appraisal methods. This in turn requires that senior administrators in a school system have a systematic knowledge of strengths, weaknesses, and applicability of different appraisal methods.

Finally, another important consideration is the cost and complexity of administering an appraisal system. Many business organizations in the past have opted for simplistic methods of performance appraisal in order to go through the motions of appraisal as quickly and cheaply as possible. While all the faults of this approach are fully recognized, it is also recognized that virtually any system which tries to satisfy all the demands on a performance appraisal system will be very complex and cumbersome to develop and administer. This will be true for a multifaceted approach to appraisal as much as for any other modern approach to appraisal. However, a type of multifaceted approach can be used which can reduce some of the complexity and administrative burden of an evaluation system.

To reduce some of the complexity, it is necessary for an organization to define or establish certain minimum acceptable standards of performance (i.e., a respectable average level of performance) and develop some simple (valid, etc.) indicators of whether a person's performance is in this range. If the primary concern of the organization is to ensure that the great majority of professionals are performing at or better than the acceptable standard of performance, then no other form of appraisal need be called into play. This would be a maintenance approach to management, much like setting a ship on a course and then making only minor corrections to the steering to remain on course. Other, more complex forms of appraisal would be used for those persons whose performance is not yet up to, or has fallen below, standard, or for those persons who wish to go through professional development to a different area of performance. Thus an array of appraisal methods would be available for use in a school system, but a "Pareto's optimum" approach could be used for the 20% of staff requiring special appraisals and simpler appraisals could be used for the 80% of staff requiring routine appraisals.

So far, the conclusions presented have assumed commonalities between the general management and the teacher evaluation literature on performance appraisal. We now turn to some of the major differences between the two sets of literature.

a) Career progression vs. organization tenure - While employees in a business organization must perform up to some level of acceptable standards to retain their jobs, there is a strong emphasis in such organizations on being able to qualify for higher level positions. The

culture of expectations in business organizations is that one must expect to have vertical career progression in the organization hierarchy. Thus the appraisal system must have a dual focus of giving feedback on current or past performance and also making assessments of potential for different kinds of work in the future.

In a public school organization, there is far less expectation of vertical career progression and instead the emphasis is on maintenance of organizational tenure or a culture of continuing to perform professional service work for the duration of a career. This is typical of professional service bureaucracies and the appraisal system consequently focuses essentially on the performance of professional service work, without any particular reference to potential for different kinds of work. Some professional bureaucracies do provide dual career ladders to a certain extent, to allow for career progression within professional service work without a person having to join the administrative hierarchy for career progression. However, even in these cases, the emphasis in performance appraisal continues to be on how well the person performs in the professional service tasks.

b) Output vs. process appraisals - Business organizations tend to have a greater use of output criteria in appraisal systems for a number of reasons. One is that many businesses have the opportunity to count or measure outputs in a variety of ways due to the nature of the production process in many businesses. Secondly, most businesses use some form of management by objectives (MBO) to manage business operations and it is a natural extension of this to use an MBO form of performance appraisal. Thirdly, there is less concern with how results are achieved (save for legal and ethical constraints) than for what is achieved in business organizations. This may be due to lack of confidence in any one style or process for best obtaining results, or due to a belief that many alternative processes are possible for achieving the same ends. In any case, the preference is to focus appraisals on results more than process.

In educational organizations, the dominant focus of appraisal systems is far more on process than on outputs. This too is consistent with professional bureaucracy theory[20] but also may be due to the difficulty (or reluctance) of measuring or counting outputs. Perhaps this is a function of public schools being indirectly financed by public tax resources, as privately financed schools seem to be much more willing to use output-based performance appraisals. The widespread use of process-based appraisals is also interesting in the light of Levin's[21] comment that so little satisfactory knowledge exists about which teaching methods do the best job of education.

c) Performance-based vs. seniority-based compensation - Consistent with use of output-focused appraisals is the use of appraisals for determining pay differentials according to performance differentials. Most business compensation systems provide some universal cost-of-living pay adjustment, but rate of progress through a job salary range or progress to a higher salary

---

20 See, for example, Mintzberg, H. The Structuring of Organizations. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

21 Levin, B. "Teacher Evaluation: A Review of Research". Educational Leadership (December 1979), pp. 240-245.

range is determined by assessment of performance. Consequently, the appraisal system in some way must be concerned with ranking or comparing one individual's work to another. Similarly, business organizations are far more concerned with justifying compensation dollars and therefore are far more prone to dismiss unsatisfactory performers. This too causes the appraisal system to have to rank-order performance evaluations in some manner.

In school organizations, the emphasis seems to be on acquiring entitlement to compensation adjustments by virtue of seniority and maintenance of adequate performance. Pay raises tend to be uniform with no performance differentials and progression through a salary range is by accumulation of years of seniority. Consequently there is no need for the performance appraisal to rank-order performance or otherwise distinguish between levels of performance except to identify serious cases of below-standard performance.

d) Non-union vs. unionized personnel - Few people at the professional service level in business organizations belong to bargaining units although they may belong to professional service associations. Consequently, their employment contracts or conditions of employment are based on uniform company personnel policies and standards plus any individual aspects of employment conditions which they may be able to negotiate. Performance appraisal will follow a similar pattern: a universal floor of common standards and practices for all, plus individuated aspects of appraisal based on the person's unique circumstances and role in the organization. In practice, there is not a lot of variation in the overall performance appraisal system used for different individuals, but the principle of individuation exists for non-union personnel.

In teaching organizations, virtually all certificated personnel are members of bargaining units and a basic union philosophy is equal rights and treatment for all members. Consequently, the nature of the appraisal system tends to be universalistic as well. This is consistent with the approach to compensation discussed in section (c) above. Individuation in teacher evaluation does exist, of course, depending upon the strengths and weaknesses of individuals and their relationships with principals, but the influence of the union is to try to maintain universal equivalence as much as possible.

e) Personal vs. non-personal accountability for results - In business organizations, an operative assumption in performance appraisal is that individuals are essentially responsible for the achievement of assignments or agreed-upon results. They are given a certain amount of resources and a certain amount of time and are expected to be the catalyst which produces the outcomes. Many results-focused systems have provision for extenuating circumstances, unforeseen events, and so on, but ultimately the individual must accept personal accountability for the results produced from their domain of responsibility.

In contrast, it has been strongly argued in the literature on teacher evaluation that individuals cannot be held fully accountable for results in the public education system because of too many uncontrollable factors or extenuating circumstances. Instead, the teacher is seen as just one of a series of actors in an educational environment which ultimately produces some form of education. Without discussing the merits of this view of

accountability, it is quite clear that the viewpoint on personal accountability for results is quite different than that found in business organizations.

f) Internal vs. public accountability - In business organizations, the conventional wisdom is that boards of directors or senior executives accept the public accountability for the performance of their organization in society. Rarely, if ever, does the public try to be involved in the internal operations of the organization except through communication with senior management. This means, in general, that accountability for individual performance within the organization is an internal matter between supervisors and subordinates. The work of individual professional service workers is less subject to public scrutiny and if the public does have some concern about the quality of this work, the matter is taken up with management rather than with the professional (an exception would be a direct fee-for-service relationship, such as with a doctor or lawyer).

In public schools, there is far greater public scrutiny of the works of individual teachers and the public is equally likely to confront the individual professional or the management of the organization. This is true even though there is no direct fee-for-service relationship between the public and the professional. The impact of this difference in concepts of accountability is that teacher appraisals have to be more sensitive to public accountability than would be true in most business organizations. This raises an aspect of complexity in teacher evaluations which would not be generally found in business organizations.

g) Whole vs. narrow performance evaluations - A "whole" job orientation in performance evaluation refers to the concept of evaluation on the complexity or multidimensional aspects of a job. If a job is relatively rich in content (i.e., extensive decision-making, autonomy, responsibility, multiskilled, long time horizons for results), then the appraisal system ought to be sufficiently multidimensional as well to be able to reflect or capture the richness of the job. If the job is relatively low in complexity or essentially unidimensional, then the appraisal system can be correspondingly unidimensional or focus on a narrow, but important aspect of the job. Despite the tendency in many business organizations to try to find a relatively simple and painless appraisal system, there is little tendency to try to focus an appraisal system on just one aspect of work. Business organizations seem to recognize that successful performance is a multidimensional phenomenon and there is an ongoing search for appraisal systems which can adequately reflect that complexity. In educational organizations, there is also a recognition of the complexity of a teacher's job, but the common practice is to focus appraisals more narrowly on specific classroom performance. Thus appraisals in schools tend to be more narrow in orientation and appraisals in businesses tend to be more whole-job oriented.

This situation does not pose a problem if the narrowly focused appraisal is concerned with a dimension of performance that is crucial to the organization or is a good indirect indicator or surrogate of the balance of the professional's work. For example, one might say that there is more to being a teacher than teaching in the classroom. However, if classroom performance is a valid representation of a person's overall performance in the school, then it

would be sufficient to rely on the relatively narrow focus of evaluating certain classroom behaviours to make inferences about other dimensions of a teacher's performance (i.e., good classroom behaviour indicates good relations with parents, indicates good planning of lessons, indicates good grading of assignments).

If these assumptions are not valid and the narrow aspect of a job which is the focus of evaluation is not a good indicator of other important dimensions of performance, then the appraisal system may be of dubious validity. For example, if classroom behaviour is independent of other important teacher behaviours, then a narrowly focused appraisal system does not really reflect the wholeness of a person's work. This situation would be exacerbated if a narrowly focused appraisal system unfortunately evaluated performance on some relatively unimportant dimension of work (i.e., neatness, punctuality) that was easy to measure, but of inconsequential validity. The question then for narrowly-focused teacher evaluations is twofold: which are the key dimensions of performance on which evaluations should focus, and are these dimensions adequate indicators of other important aspects of a teacher's work? The alternative is to develop multidimensional appraisal systems similar to business organizations which try to capture the complexity of the work of a professional.

In summary, all of the above points suggest a series of cultural or philosophical differences which exist between performance appraisal in business and teacher evaluation. These are differences which have evolved over time and reflect the different nature of operations in the two types of organizations. However, the differences are not immutable and some possibilities for convergence of practices can exist. These convergences should not be sought merely for the sake of convergence. In keeping with the theme of conclusions from this report, appraisal practices from the business world should only be borrowed when they make sense to the needs and priorities of a particular school organization or for a particular segment of the certificated staff. Similarly, historical precedents in teacher evaluation should only be retained or perpetuated if they continue to make sense or meet the needs of a particular school system. Ultimately such judgements will have to be made by persons involved in each particular school system.

3. FIELD SITE RESULTS

3.1 INTRODUCTION

The purpose of this project was to conduct a series of field study site visits to organizations which are analogous to public education organizations in that large numbers of professionals are employed and formal appraisals are conducted of professional service performance. The field site study was organized into two parts: a study of ten analogous organizations in Ontario which could provide useful insights into the appraisal of professionals, and a study of selected public school boards in the United States which have acquired a reputation for being very progressive in their approach to appraisals of educational staff.

In general, the study of analogous organizations focused upon the following issues:

i) to identify the appraisal methodologies and the uses of appraisal in the various types of organizations;
ii) to identify major differences (if any) in approaches to appraisal which seem to be required by the different organizational types;
iii) to identify differences in approaches to appraisal in different functions of the organization and in different levels of the hierarchy;
iv) to identify whether process or output approaches to appraisals are typically used in some functions and organizations and not in others;
v) to identify various approaches to developing criteria and gathering data on the work of professionals; and
vi) to identify linkages between the appraisal process and other parts of human resource management in the organization, such as compensation systems, training and development, career planning, and supervision.

3.1.1 The Business Group

The Business group was identified by consultation with colleagues at the Personnel Association of Ontario and the Compensation Research Centre of the Conference Board in Canada. The criteria for sample selection were employment of a high proportion of professional service employees and a good reputation in the personnel community for general personnel management. From this it was assumed that such organizations would also be examples of good practices in performance appraisal since performance appraisal is a foundation for so many other aspects of personnel management. The identity of each organization has been kept confidential, partly to encourage free-flowing, frank discussions about performance appraisal and also because some sites requested confidentiality as a condition for participating in the study. Out of a list of ten recommended sites, five organizations agreed to participate when first approached and the remainder of the list was not used.

3.1.2 The Colleges Group

Colleges of Applied Arts and Technology (CAATS) were selected for this study because they obviously employ a large number of teaching professionals and are located in the Ontario educational environment. They are different from schools in public education in that they are post-secondary, tuition-based schools and many of their instructors are not formally trained or educated as teachers. However, the colleges can be a useful sample to study for performance appraisal practices because they have to be more directly accountable to their students for a good education (tuition fees) and to the local employment community. Many colleges use advisory boards of local employers and prepare students for Provincial trade certificates or licences. Consequently, the performance of a college in producing qualified students can be closely scrutinized in many ways. It follows from this that many of the colleges have devoted careful attention to maintenance and upgrading of teaching quality.

The colleges are required by their collective agreement with teachers to provide appraisals of teaching effectiveness once a year. The method of appraisal and the dimensions of performance are not specified and it is up to each individual college to devise its own system. Practices ranged from lip service to appraisal, to different approaches within different departments in the same college, to integrated systems with quite serious attention paid to appraisals.

The sample for this study was identified by a study group of CAAT personnel directors for the purpose of sharing information on performance appraisal practices at each college with the eventual hope of developing some common (voluntary) standards across the Province. The study group of personnel directors was able to identify five colleges as being the most advanced or at least the most seriously concerned with developing effective appraisal systems. These five colleges are identical in types of professionals employed (lecturers, teaching masters, area co-ordinators, and department heads). Although these colleges were not selected to be a random or representative sample, there is no particular geographic region, program specialty, or size factor in common among the group studied, so they represent a reasonably good cross-section of the colleges in the Ontario System.

3.1.3 The U.S. Schools Group

The third type of field site sample was a group of United States public school systems which had been identified through the literature review as being front-runners in either innovative or serious approaches to teacher evaluation. Although the U.S. Schools group obviously operate in different education and employment environments than Ontario, they do offer the opportunity to examine a number of state-of-the-art approaches to teacher evaluation which might be applicable or relevant to Ontario schools. Secondly, many of the States have state-wide, mandatory teacher evaluation systems which serve as a baseline for evaluation practices in local school districts.

The U.S. Schools group consisted of seven public school systems as well as the teacher evaluation service of Educational Testing Services (ETS) of Princeton, New Jersey. The seven school systems are:

i) The Chicago Board of Education, Chicago, Illinois;
ii) Montgomery County Board of Education, Rockville, Maryland;
iii) High Point Board of Education, High Point, North Carolina;
iv) State of Connecticut Department of Education, Hartford, Connecticut;
v) State of Georgia Department of Education, Atlanta, Georgia;
vi) Oakland Unified School District, Oakland, California;
vii) Orange County Board of Education, Orlando, Florida.

There is no particular geographic or urban pattern to the U.S. Schools group. Sites range from large cities (Chicago, Oakland) to smaller cities (High Point, Orlando), from northern industrial states (Connecticut, Maryland) to Southern (Georgia, Florida). It was hoped that such a distribution of sample sites would provide a diverse range of approaches and concerns in teacher evaluation systems and not be limited to the special problems of any one geographic or socio-economic area.

3.2 PROCEDURE

Each of the prospective Business and College sites in Ontario was contacted by telephone to arrange interviews with the director of personnel or the chief administrator of the performance evaluation system. Subsequent interviews were then arranged with department heads or supervisors who actually conducted appraisals to obtain their perspectives on their experiences with the system.

Occasional interviews were also possible with professionals themselves on their experiences with being appraised by these systems. No attempt was made to try to gather widespread evidence from appraisees on satisfaction with a system. This would have required extensive surveys with numerous samples of respondents and the logistics of trying to conduct a valid and reliable study of that sort was beyond the resources of the field study project. Instead, satisfaction with a system had to be inferred from three admittedly weak indicators: testimony on satisfaction from the interviews, numbers of revisions felt to be necessary to a system, and record of grievances, appeals, arbitrations, and lawsuits arising from appraisals.

Site interviews with the Business and Colleges group were conducted during a period from August to December, 1982. Consequently, the information gathered reflects practices and systems in effect or planned at that point in time. Prospective sites in the U.S. Schools group were contacted by telephone by one of the research team to arrange a site interview. All sites approached were willing to participate in the study and site visits were done during the period September 1982 to May 1983. Interviews were conducted with the senior administrator responsible for the teaching evaluation system and occasionally with specialists responsible for design, implementation, monitoring, or evaluation of the system. Interviews at all school sites were also conducted with supervisors (superintendents, inspectors,

principals) who used the appraisal system and, occasionally, interviews were done with teachers subject to the evaluation system. The same indicators for relative satisfaction with a system were used for the U.S. Schools sites as were used for the Business and Colleges groups.

3.3 RESULTS

Results will be reported for the following topic areas:

a) Initial Design and Implementation of the System;
b) Current Objectives and Purposes of Appraisal;
c) Appraisal Design and Techniques;
d) Criteria or Standards for Appraisal;
e) Sources of Data on Performance;
f) Training and Preparation for Apraisal;
g) Special Cases of Appraisal and Organizational Justice;
h) Evaluation and Maintenance of the Appraisal System;
i) Trends and Outlook for the Appraisal Systems.

3.3.1 Initial Design and Implementation of the System

A comparison between the three types of samples can be made by looking at reasons for developing the appraisal system and the method by which the system was designed and implemented. Table 1 summarizes these comparisons. As can be seen from this table, there is a fairly clear distinction between the Business group and the two groups of educational organizations in the reasons for developing an appraisal system. The Business group reasons, in order of importance, were (1) improved organization performance, (2) improved personnel administration, and (3) improved quality of professional service. In both types of educational organizations, the reasons were (1) mandatory requirements for teacher appraisal, (2) improved quality of professional service, and (3) improved personnel administration. These priorities may differ in individual organizations but the order of priorities represents the patterns for the sample as a whole.

In terms of design and implementation of a system, the top management approach to design is predominant in the Business group while the joint or collaborative approach is favoured in the educational organizations. It is possible that a relationship exists between reasons for developing a system and method of developing a system. The business organizations were primarily concerned about improving organization performance and used a top management approach to this end. The educational organizations had a high priority on quality of professional service and used a collaborative approach to the design of an appraisal system. However, this pattern is less clear for the Schools group since three of them used a top management approach and were also concerned about quality of education. Two of the schools with the top management approach (Chicago, Oakland) were in a politically demanding situation and may have behaved similarly to the Business group under pressure to produce system-wide improvements in a short period of time.

Table 1

Initial Design and Implementation of the Appraisal System

		Business Group (5)	Colleges Group (5)	U.S. Schools Group (7)
A.	Impetus priorities for developing a new system	(order of importance)		
1.	Improve organization performance	1	-	3
2.	Mandatory requirement	-	1	1
3.	Improve quality of professional service	3	2	2
4.	Improve personnel administration	2	3	-
B.	Methods of development of the appraisal system		(frequencies)	
1.	Top management design and development	4	2	3
2.	Joint committee design and development	1	3	4

3.3.2 Current Objectives and Purposes of Appraisal

Table 2 provides a summary comparison between the three samples on a number of general purposes served by the appraisal systems. Some of the purposes may be implicit in a number of organizations (i.e., to improve student learning in schools), but are only listed in this table if they were explicitly articulated at the sites. An example of this concept is the use of appraisal information for summative evaluation and personnel decisions, such as compensation, tenure, and promotions. The majority of the Business group are explicit in stating that such information has a direct linkage to personnel decisions, but the majority of schools and colleges refer to such information as "possible inputs" to personnel decisions. This means that appraisal information may be called into play if performance problems arise or are identified in some way, or that personnel decisions such as compensation progression and seniority accumulation occur automatically as long as a person's performance is "satisfactory". However, all of the sites studied admitted that appraisal information has to be the basis for taking action and justifying such actions if performance problems occur. In these situations there is a direct link between formative performance appraisal and summative personnel decisions in all organizations.

Table 3 provides a summary comparison of more specific, routine purposes of appraisal systems between the three samples. The major differences appear in aspects which are more characteristic of business organizations (i.e., performance pay, promotion potential). Another major difference is that identification of training needs is a common purpose of appraisals in business organizations. The purpose should be just as prevalent in the U.S. Schools group given the emphasis on professional development found in most of these sites and the absence of this purpose suggests that the emphasis on professional development in the Schools group is not consistently followed through for individual teachers. Instead, the general approach to professional development is feedback, coaching, and guidance; general school board professional development activities; and individual initiatives in training activities.

3.3.3 Appraisal Design and Techniques

Table 4 provides a summary of the information in the three samples. From Table 4 it can be seen that all three types of samples have a hierarchical difference in approach to appraisals. Without exception, administrators and managers use a goal-setting and action-planning style of management and their performance is evaluated either on the completion of their plans or the results they achieve. In general, this goal setting and planning by administrators is done in the context of organization goals and plans. Between divisions or departments, a uniform format of appraisals is used by sites in both the Business and U.S. Schools groups. Only in the colleges were some examples found of different appraisal practices being used for jobs at an equivalent level in different departments in the same organization. While it can be argued that the nature of work may be sufficiently different between departments to justify different appraisal methods, it should be recognized that such practices are open to claims of inequity in the organization. The other difference in general is that most of the U.S.

Table 2

Current Objectives and Purposes of Appraisal Identified for the Organizations in the Study

	Objectives/purposes	Business(5) (number who claim these objectives)	Colleges(5)	U.S. Schools(7)
1.	Improve student/client learning	1	1	1
2.	Improve quality of service/teaching	2	5	3
3.	Improve organization management	5	2	2
4.	Improve professional development	5	3	4
5.	Provide a direct link to personnel decisions	4	-	-
6.	Provide inputs when required for personnel decisions	1	5	3

Table 3

Specific Purposes of Appraisals

Specific Purposes	Business Group (5) (number who claim these	College Group (5) purposes)	U.S. Schools (7)
1. Seniority maintenance	3	5	7
2. Eligibility for performance pay	3	0	0
3. Identify training needs	5	5	0
4. Discover performance problems	1	0	0
5. Document performance problems	5	5	7
6. Assess promotion potential	5*	0	0

*Note: The Business group distinguish between assessing competence as a professional and assessing potential for promotion. A good record of performance as a professional is a prerequisite to eligibility for promotion. Special performance appraisal techniques (i.e., assessment centres, trial runs as supervisors, acting project coordinator roles, etc.) are used for appraisal of promotion potential.

Table 4

Appraisal Design and Techniques

Types of Design Techniques	Businesses (5)	Colleges (5)	U.S. Schools (7)
a) 1. Differences in hierarchy	5	5	5
2. Differences across divisions	1	3	0
b) 1. Focus on certification/competency	1	5	6
2. Focus on work processes	1	5	7
c) 1. Focus on action plans	4	4	3
2. Focus on results	4	1	3
d) Formats of appraisals			
1. Summary essays, memos, letters, etc.	3	5	3
2. Tests, rating scales	2	1	4
3. Student/client evaluations	1	5	2
4. Checklists of abilities, skills	4	4	3
5. Forced-choice comparisons	0	0	0
6. Critical incidents record	1	0	1
7. Ranking	2	0	0
8. Semi-structured observations	0	3	6
9. Unstructured observations	0	0	1
10. BARS, BORS	0	0	0
11. MBO - action plans	4	4	3
12. MBO - results	4	1	3
13. Self-evaluations	2	3	3
14. Peer evaluations	1	0	0
15. Numerical summaries of performance	3	0	2
e) Interim feedback	5	5	4

Schools had a strong emphasis on certification or competency evaluation for entry level jobs which was not found in either the colleges or the business organizations (with the exception of one accounting firm). This means that many of the U.S. Schools group had to verify competency after a person was employed while the other organizations assumed competency at the time of hiring and instead concentrated on verifying level of performance.

In terms of the focus of appraisals, both the colleges and the U.S. Schools had a pronounced focus on process in their evaluation systems, while the Business group had a pronounced focus on results. Some of the U.S. Schools had a hybrid approach by focusing on process and results but this was less frequent in the other two samples.

3.3.4 Criteria or Standards for Appraisal

Table 5 summarizes the information in this section for the three samples of organizations. There is a roughly equal tendency among all three groups to have general job descriptions for professionals and little tendency to have very detailed job descriptions. However, two of the U.S. Schools do not have job descriptions at all for teachers which is a significant difference from the Business and Colleges groups.

All three samples have established licence, degree, or certificate requirements as minimum basic standards for qualifying for a position but only some of the U.S. Schools group do their own certificate testing. The practice of defining specific competencies and professional service behaviours is far less common in the Business group than in either the College or U.S. Schools groups. The specific competencies and behaviours in the colleges are defined in the student evaluation forms and in the U.S. schools by competency criteria.

It is fairly common in all three samples to have some specification of professional service behaviours or styles as part of the appraisal criteria. In the Business group, these were agreed upon methods or corporate styles for professional service work. In the colleges and schools, these were criteria specifications for the competencies required. (The criteria specifications are often written in behavioural terms or guidelines.) Only one of the Business group (accounting) also used specific process standards and criteria for appraisal, while these were far more commonly used in the colleges and schools. Typical criteria consisted of teachers' lesson plans and teaching activities which then formed the basis for observations by principals, or other supervisory personnel. This approach of using rating scale or semi-structured observations of a professional's work by a supervisor was practically non-existent in the Business group.

Instead, the Business group was consistent in the use of results specifications for appraisal, both in terms of action plans and results specifications of outcomes. The colleges and some of the schools were active in specifying plans for objectives, but rarely were results themselves specified. This means that completion of the plans themselves were accepted as results for evaluation purposes rather than verification of whether the plans actually achieved the intended objectives.

Table 5

Comparison of Standards and Criteria of Performance

Types of Standards and Criteria	Business (5)	Colleges (5)	U.S. Schools (7)
a) Job descriptions - general	4	5	3
- detailed	1	0	2
b) Minimum standards specifications			
- licence, degree or certificate only	5	5	7
- specific competencies; behaviours	2	5	4
c) Process standards specification			
- professional service behaviours, styles	4	3	6
- specific process plans, activities	1	4	7
d) Results specifications - plans	5	4	4
- outcomes	5	1	1
e) Professional development - plans	5	5	5
- outcomes	3	0	1
f) Contribution to organization plans	5	3	2
g) Compatibility with organization culture			
- explicit	3	2	1
- implicit	2	0	3
h) Appraisal orientation - whole	5	4	2
- narrow	0	1	5

Professional development plans were used in all of the businesses and colleges and in many of the schools. The businesses tended to look more closely at the results of the development plans and therefore these were used more seriously for appraisal purposes than in either colleges or schools. In all cases, development plans and results were examined very carefully in special review situations of problem cases in performance. Another pronounced difference between performance criteria in business and the two types of educational organizations was the emphasis in the Business group on contribution to organization plans and compatibility with organization culture. These criteria are consistent with the greater use of goals-based management and appraisal systems in the Business group, and the effect is to have a wider or more of a whole-job approach to appraisals in the Business group. Despite the widespread use of MBO at the administrative levels of both the colleges and schools, instructors and teachers are evaluated by a much narrower range of criteria and consequently there is a narrower operational concept of the work of the professional.

3.3.5 Sources of Data on Performance

Table 6 summarizes the results for each of the three groups of organizations. The Business and College groups have much more frequent formal appraisals than the U.S. Schools although the average practice across all three samples is to have a formal annual appraisal. The use of bi-annual and tri-annual formal appraisals in the U.S. Schools sites was generally reserved for senior, tenured teachers in whom principals had a lot of confidence. Provision was made in all organizations for shorter term formal appraisals where problems seemed to exist. However, the schools with up to three years between formal appraisals do run a greater risk of performance getting out of hand or being difficult to correct, by the time that it is concluded from formal appraisals that a performance problem exists.

Despite frequent recommendations in the literature, the use of formal pre-observation conferences was not common in either the colleges or the schools. Many of the educational sites had such conferences as an optional practice, but the exercise of the option was not especially frequent. This seems to be a shortcoming in appraisal practices where so much depends on classroom observations by principals or department heads. As was stated previously, many supervisors follow standard observation forms or rating scales in conducting the observations so perhaps it can be argued that there is not much to be discussed in a pre-observation session. It was also claimed at some sites that the sheer magnitude of the number of required classroom visits prohibited much opportunity for discussion beforehand (and afterwards). However, for those systems which require their teachers to develop teaching strategies, lesson plans, and so on, it would seem necessary for a principal to have a pre-observation conference in order to establish what the principal was supposed to be looking for in the classroom visit.

The Business group do not use any form of formal pre-observation conferences but instead rely on the annual planning and goal setting discussions to accomplish the same purpose. These discussions also alert the supervisor as to what to expect and look for in a subordinate's performance, but the time period could be for up to a year afterwards and the observations less formal.

Table 6

Sources of Performance Appraisal Information

Sources		Business (5)	Colleges (5)	U.S. Schools (7)
a) Formal Appraisals	i) semi-annual	4	2	-
	ii) annual	1	3	4
	iii) two years	-	-	1
	iv) three years	-	-	2
b) Formal pre-observation conference		0	2(+3 optional)	2(+2 optional)
c) Formal post-observation conference		5	5	7
d) Formal direct observation - supervisors		1	5	7
- experts		0	0	1
e) Formal self-reports of performance		3	3	2
f) Formal reports by clients/students		2	5	2
g) Formal reports by subordinates		0	2	1
h) Formal measures of client/student performance		2	2	2
i) Supervisor monitoring and informal observations by supervisor		5	5	7

Post-observation conferences are used by all of the organizations in the study. The difference between the educational organizations and businesses is that the conferences in colleges and schools take place very soon after the formal observation of performance. In the Business group, the post-observation conference is the semi-annual or annual appraisal conference which serves as the summative evaluation of the overall performance for the period. In terms of learning theory, the practice in educational organizations is more effective because feedback occurs shortly after the observed performance. The effect is similar to a formative appraisal even though the intent is to be summative. On the other hand, the practice in business is to try to take a whole job perspective for the whole period of performance and these appraisals would not be as susceptible to recency bias as might a teacher appraisal.

Formal, direct observation by supervisors was the most common source of appraisal information in both of the educational groups and was almost nonexistent in the Business group. This means that educational organizations rely extensively on observations by department heads or principals as the most basic source of appraisal information. Although they are aided by checklists, rating scales, criteria references, and observation guides, the performance appraisal of a teacher ultimately boils down to what a supervisor observes in a classroom. Obviously this places a considerable burden on such supervisors to be competent observers so that the data derived from this source are valid and reliable. The other obvious problem with formal direct observations by supervisors is the short time period, the relative artificiality, and the narrow focus of such an approach.

To compensate for these limitations, most organizations tend to use some supplementary sources of information about a person's performance. A few in each sample used self-reports on performance by subordinates/teachers, and some used reports or evaluations by clients or students. Student evaluations of courses and teachers were used in all of the colleges, due to the high emphasis by CAATs on good quality teaching (and the implicit assumption that students could provide valid information about teacher performance). The colleges were also careful to try to distinguish between evaluation of a course (which is akin to program evaluation) and evaluation of an instructor. Formal reports or ratings by subordinates and teachers were used in a few cases in colleges and schools for the appraisal of principals or department heads. There did not seem to be any problems with acceptance of this procedure.

Formal measures of client/student performance were used in a few sites in each of the three types of organizations. In all cases, this information was used more for program evaluation or evaluation of an organization than for individual professionals within the organization. Of course program and organization evaluation can be used for evaluation of the responsible administrator because it is consistent with the use of a results-based appraisal system for administrators and managers. No evidence was found of student performance criteria being used for the evaluation of individual teachers or instructors, although this could be done if some individual professional were found to be responsible for the negative (or positive) results of such program evaluations.

Finally, the most common source of appraisal information in all three samples was monitoring and informal observations by supervisors. As stated previously, this sort of information is inevitably acquired as a result of the routine interactions between people working together in an organization. It is part of the supervisor's role in an organization to convey expectations about performance, to give feedback on whether performance is consistent with expectations, and to provide help and counsel where required, or support and recognition where merited. In the course of these duties and responsibilities, the supervisor cannot help acquiring a cumulative series of impressionistic judgements about a subordinate's performance. In addition, supervisors typically find themselves to be recipients of a variety of unsolicited information about performance in the organization. Unfortunately these are often complaints, but occasionally can be compliments. As well, if there happens to be a sufficiently well established culture peculiar to the organization, supervisors can find themselves in the unofficial (and unconscious) role of socialization agent for the organization.

All of these factors combine to place a burden of responsibility on supervisors similar to that described for direct observations of performance. However, the supervisor is not usually aided by a clear set of guidelines, competencies, and criteria for making informal observations and impressionistic judgements. Supervisory monitoring and informal observations provide a rich source of data about a person's work performance and help to capture the wholeness of a job better than many of the formal sources of data. The challenge is to minimize the likelihood of bias or limited validity and reliability in such data. The existence of this challenge highlights the need for training in performance evaluation for supervisors since they seem to be the key link in virtually any form of performance appraisal found in this study.

3.3.6 Training and Preparation for Appraisal

Table 7 provides the summary results for this section. The results previously reported in this study have established that the supervisor plays a pivotal role in evaluation of performance. The supervisor is the primary communicator/negotiator of expected standards of performance, the primary observer or data gatherer about performance, and finally, the primary deliverer of feedback to subordinates about how their performance is perceived and evaluated. In recognition of the importance of these supervisory roles, it should follow that organizations provide considerable support and training for supervisors on how to do performance evaluation.

All of the sites in the three samples in the study had extensive manuals, procedures, and policy guidelines on performance evaluation available for supervisors. Support was also provided by the respective personnel departments in the form of administration of the system to ensure that appraisals were done on time, and according to procedures. (Some might perceive this as pestering by the personnel department rather than support.) Direct support was also provided in most business, colleges, and schools by having supervisors discuss

Table 7

Training and Preparation for Appraisals

Types of Training and Preparation	Business Group (5)	College Group (5)	U.S. Schools Group (7)
1. Formal support systems	5	5	7
2. Informal support systems	5	5	7
3. Formal training for supervisors	2	1	2 (+2 optional)
4. Informal learning for supervisors	5	5	7
5. Formal orientation for subordinates	5	5	7
6. Informal learning for subordinates	5	5	7

Note: Organizations could obviously use formal and informal practices together.

appraisals with someone from a more senior level of management. This secondary level of discussion accomplished two purposes. Senior managers would be kept in touch with general levels of performance of a wide range of personnel and evaluations of personnel between divisions or departments could be monitored. This procedure was important for the early identification of performance problems and for ensuring equity of treatment across divisions in a particular organization.

Formal training in conducting performance appraisals was quite different between the samples. Only two of the Business group and one of the College group had formal training programs for supervisors in how to conduct appraisals. It appeared that most supervisors were supposed to learn how to do appraisals by informal methods. On the other hand, four of the U.S. Schools had some form of appraisal training although it was optional in two sites and subject to budgetary cutbacks. The three other U.S. School sites relied on detailed procedural manuals and informal learning.

All the U.S. Schools, and some of the businesses and colleges, had formal introduction and briefing sessions for administrators and supervisors when new performance evaluation programs were introduced. However, the level of maintenance training or training for subsequent new supervisors rapidly dropped off. These findings raise some doubts about the thoroughness of the commitment to effective performance appraisal programs which exists at senior levels of management, or suggests a strong belief in the efficacy of informal learning about a system.

Finally there is also the question of training of subordinates in the appraisal process. Conceivably, subordinates should be equally well trained in how the system is supposed to work and what is to be expected of them in the appraisal process. In general, the three groups of organizations did try to convey to subordinates the sorts of performance expectations that would be used for appraisals and the formal methods by which information on performance would be gathered.

This orientation or training allows subordinates to be somewhat self-managing in that they know to some degree what is expected of them and how these expectations will be evaluated. This is consistent with the notion that professionals should work in a fairly independent manner, largely self-responsible for the guidance of their conduct. What is missing however, is training in how subordinates should behave in the processes of evaluation; i.e., the setting of goals and action plans, the development of performance criteria, the preparation of self-evaluation and the discussion of development plans. How to perform well in these sorts of activities did not seem to be the subject of formal training in any of the study sites and it seems to be something that has to be learned informally through interaction with the supervisor. Thus one more role burden is suggested for supervisors: training and counselling of subordinates in how to go through the appraisal and evaluation process.

3.3.7 Special Cases and Organization Justice

The detailed arrangements for a special review of performance problem situations and for provision of organization justice vary widely from site to site, but the general principles are much the same. The elaboration of procedures is normally much greater in the unionized colleges and schools than in the business organizations. However, the personnel departments in some businesses have developed equally detailed procedures in the absence of a union. In addition, legal trends are forcing all organizations to be much more careful and thorough in dealing with problem situations than in the past. The results in this section will be discussed in two parts: probationary employees and established employees.

a) Probationary employees: Almost all of the organizations in this study provided a fairly stable work environment with job security or quasi- or formal tenure available to persons with satisfactory performance evaluations. For most of the organizations in this study, a psychological or implied contract exists between the employer and subordinates that once the probationary period of employment is satisfied, then job security is essentially assured.

Along with this implied contract of quasi- or formal organization tenure, is also the implied commitment of the organization to using internal sources of human resources for most higher level staffing needs (i.e., for all supervisory and managerial jobs but excluding specialists and perhaps very senior executives.) Thus it is incumbent on the organization to be very careful about who is recruited to work for the organization and, especially, who is allowed to stay after probationary employment.

Six of the U.S. Schools have quite detailed evaluations of probationary teachers. The content of the evaluations focus on both subject area knowledge and competency on whatever criteria for good teachers have been established by the school jurisdiction. Formally or informally, the school jurisdiction creates a certification process in addition to the teaching certificates one might earn from a teachers' college. The probationary period of employment lasts until the completion of the certification process.

Once the person has been assessed as qualified or competent (within the allowed time limits), then the teacher is in effect licensed to practice in that school jurisdiction. The seventh U.S. School site is a school board which accepts teacher certificates issued by the State for admission to probationary employment and then looks to see if a teacher receives satisfactory performance ratings during the probationary period. There are no special criteria, detailed appraisals, or attempts to create local certification for probationary teachers in this seventh site.

Periods of probationary employment may last from one to two years and opportunities for repeat tests or trials may be provided, but the ultimate result is the same. If a person fails to qualify in the mandatory probationary period, then the teacher is not allowed further employment in that school jurisdiction. Opportunities to receive counselling and coaching are generally available for teachers who are having difficulty qualifying, but the conditions of

dismissal or acceptance are clear and objective. Thus there are no legalistic disputes concerning dismissals as a result of failing to qualify during the probation period.

The colleges also pay very careful attention to instructor performance during probation because of the concern for instructor quality, and the fact that few instructors have received prior training as teachers. Some of the colleges have established sets of required instructor competencies similar to the U.S. Schools, but there is no attempt to do formal local certification or licensing of instructors. Instead, the probationary instructor has to receive satisfactory course and student evaluations during the probationary period and if successful, is hired as a member of the "permanent staff".

In two of the Business group the minimum standards and criteria for probationary employees were well established and failure to qualify was treated similarly to the schools and colleges. The other three businesses had less definite performance expectations for probationary employees and consequently used more individualistic judgements about whether a person qualified for "permanent" status. Generally, if probationary individuals received satisfactory evaluations in the same appraisal process used for all other professionals, they were eventually considered to have passed from probationary to permanent status. Informally, a one year probationary period was used with opportunities for qualifying possibly extending for another one to two years. In summary, the U.S. Schools are the most formalistic and careful in probationary performance appraisals, the colleges are less formalistic but equally careful, and the businesses less formalistic and less careful.

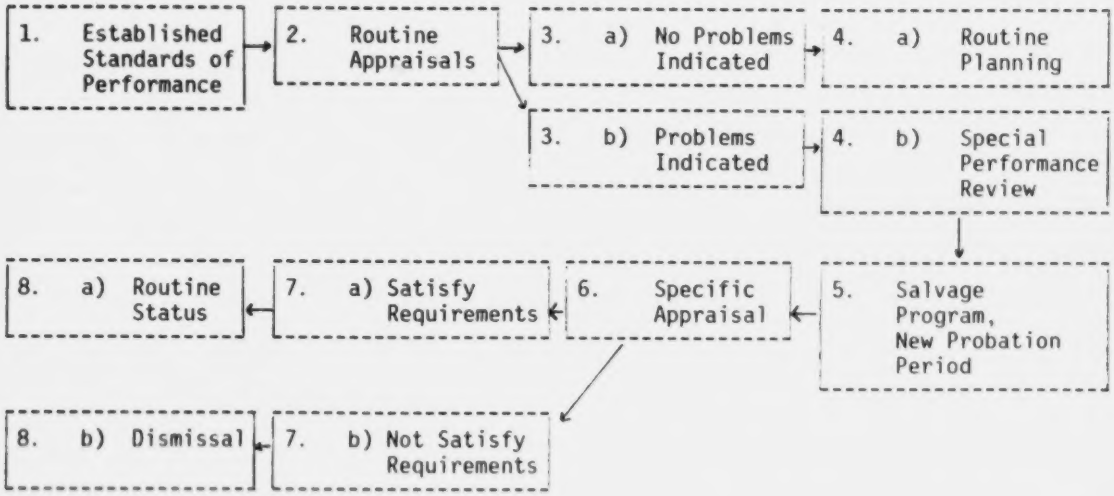
b) Established employees: This term is used to describe employees who have satisfactorily completed probation and are now regarded as essentially permanent employees of the organization. Their tenure may be formal or informal and as long as they continue to receive some satisfactory minimum rating in performance evaluations, they continue to accumulate seniority in the organization. If established employees start to receive less favourable performance evaluations, then a special review process or renewed probationary status is triggered into action. This was the case in all of the organizations in this study. The detailed procedures of the special review process varied from site to site but the essential features of the systems were the same.

The essential features of the special review process are as shown in Figure 1, which outlines a model of a performance-monitoring system. The performance-monitoring system is easier to manage if performance standards have some clearly established minimum levels and criteria. If the minimum acceptable levels of performance are nebulous, then it is much more difficult to identify performance problems. Most of the schools and colleges were more careful to establish minimum required standards and two of the businesses did so as well. Those who left these criteria unspecified were subject to contentious debates when attempts were made to confront performance problems.

Salvage programs could go by many names but they all involved supervisor-subordinate discussions of indicators of unsatisfactory performance, analyses of causes, coaching and/or corrective plans, and subsequent specific appraisals on the results of the plans. Eventually

Figure 1

Example of a Performance Monitoring System with Special Performance Review


1. Established Standards of Performance
2. Routine Appraisals
3. a) No Problems Indicated
4. a) Routine Planning
3. b) Problems Indicated
4. b) Special Performance Review
5. Salvage Program, New Probation Period
6. Specific Appraisal
7. a) Satisfy Requirements
8. a) Routine Status
7. b) Not Satisfy Requirements
8. b) Dismissal

a decision has to be made as to whether inadequate performance is correctable or has been corrected, or whether the employee is judged to be incompetent for the job requirements. If the level of performance persists at an inadequate level despite the best efforts of a salvage program, then the evidence points to incompetence as the cause.

Evidence for demonstrated incompetence (provided by the monitoring and a special review process) has generally been accepted by the courts and arbitration systems as sufficient grounds for dismissal for cause. Other options used for dealing with incompetence were transfer of the individual to a less sensitive or demanding job, redesign of a job so as not to require the deficient competencies, and to have frank, career planning discussions with problem employees about whether they really belong in the particular profession or would be better suited to a less demanding organization. Due process, fair treatment, and rigorous documentation of all steps and procedures in special performance review cases are necessary for the protection of both parties.

All the sites in the study have developed systematic and carefully documented procedures for whenever serious performance problems occur among established employees. In practice, the actual use of such procedures is comparatively rare. One reason is that rigorous screening at the probationary level removes most potential performance problems at the outset. Another reason is that the special performance review process is time-consuming and costly and the organizations have to go through some very careful consideration of whether a performance situation is sufficiently serious to warrant the procedure. Finally, counselling sessions with inadequate performers often result in voluntary withdrawal from the organization. Employees with a performance problem may also not wish to go through the complications and tensions of the special review and salvage program process and, instead, will try to seek alternative employment once they know where they stand in the present organization.

One difference which was found between the Business group and the educational organizations concerns employees who consistently receive performance evaluations at minimum satisfactory levels. Technically these employees satisfy the conditions of employment, but they are marginal or mediocre employees in the eyes of their employers. The schools are basically obligated to leave these teachers alone if they are minimally satisfactory because education organizations have less discretionary latitude for dismissing marginal performers than do business organizations. The colleges in this study try to avoid a situation of marginal performers by requiring a high level of teaching performance as a minimum acceptable standard during the probationary period. If what might be called "teaching excellence" is demonstrated during probation, the colleges hope that long-run performance will continue at an equivalent level.

Most of the Business group have the philosophy of trying to require excellence as a minimum standard of performance. Persons who consistently receive marginal, barely acceptable, or below average evaluations are not assured of job security in these organizations. Professionals with this sort of record receive the lowest possible salary increments, do not receive promotions, and have to participate in counselling discussions about their suitability to work in that organization.

3.3.8 Evaluation and Maintenance of the Appraisal System

The evaluation of an appraisal system is essentially a matter of program evaluation, that is, an assessment of the outputs and consequences of a program in comparison to the goals established for the program. For the organizations in this study, it is safe to say that none are conducting formal program evaluations of their performance appraisal systems. This is partly due to recognition of the complexity of what a formal program evaluation requires, but also due to the limitations of resources within the organizations. Most resources are devoted to implementation and administration of the appraisal systems and very little is left over for other purposes. Consequently, any form of program evaluation which does occur is in the form of impressionistic evidence and indirect indicators.

Among the Business group, three of the appraisal systems were felt to be too recently launched for any sort of program evaluation to be done. These systems are operating on management's faith or conviction that the approach is appropriate for their organizations. No revisions are yet felt to be necessary and no record of grievances has yet been established. Indirect indicators such as paperwork burden or sense of frustration with the systems are also not yet apparent. The other two appraisal systems were also launched under the strong conviction by management of the appropriateness of the system and their experience with their systems (both MBO) has not challenged these convictions. There is no particular record of grievances, no frustrations with paperwork, or other indirect indicators of problems. Consequently, the conclusion from the impressionistic evidence is that the appraisal systems in the Business group are operating satisfactorily and no significant changes are planned.

A similar situation prevails in the College group, although a few differences are apparent compared to the Business group. Since some of the colleges used joint task forces to develop their appraisal systems, these systems were launched with the conviction of both managers and instructors that the systems were appropriate. The operations of the appraisal systems are fairly new in all the colleges but the experience so far has not suggested the need for any significant changes. Grievances and arbitrations have occurred in a few situations. These have been concerned with attempts by colleges to dismiss some established instructors for continued inadequate performance. These are very few in number in comparison to the large number of instructors employed at the five colleges. In terms of frustrations with appraisal systems, there are more complaints about paperwork in the colleges than in the businesses. This is interpreted as being due to the greater care paid to probationary evaluation; the greater use of direct supervisor observations and client reports; and the union membership of the instructors. The recent launching of a centralized MBO planning system at the colleges also contributed to the perception of increased paperwork.

Two of the U.S. Schools had a form of program evaluation in that student learning and student achievement test scores were used as criteria for the validation of teacher competencies. Since the teacher appraisal systems were founded on the assessment of competencies, then student performance could act as an indicator of the effectiveness of the appraisal system as well. This type of program evaluation was most extensively developed in the State of Georgia where student achievement records were used both to validate the required

competencies and the effectiveness of the teacher certification program. Otherwise, the U.S. Schools followed much the same impressionistic evidence approach to program evaluation as did the colleges and businesses.

Overall, the impressions received from the sites in the study reveal few formal attempts to evaluate the success of the performance appraisal systems and little evidence of any problems with the systems. If problems should begin to develop, then they will probably be recognized only after indirect evidence mounts to a level considered to be intolerable.

3.3.9 Trends and Outlook for the Appraisal Systems

As might be gathered from section 3.3.8, the Business, College, and U.S. Schools groups are basically satisfied with their current systems and may only introduce minor changes to fine-tune or complete stages of introduction of their systems.

All three types of organizations will continue to require employees to maintain a satisfactory minimum level of performance to retain their seniority, although the definition and criteria for "satisfactory minimum" will vary from one organization to another. Each will conduct routine appraisals using various approaches to verify this level of performance requirement. Each will use some form of goal-setting and action-planning in addition to routine appraisals. In the businesses, goal setting and planning will be mostly concentrated on contribution to organization goals with some aspects of professional development planning. In the colleges, there will almost be a balance between professional development plans, contribution to organization goals, and teaching plans. For teachers in schools, goal setting and planning will be dominated by teaching plans, with some professional development plans, and little planning for contribution to organization goals. This latter pattern is somewhat unfortunate for teachers because it implies a role for them as teaching mechanics rather than professional educators. Instructors in the colleges have a more professional role as educators with a more holistic approach to goal setting (and a more holistic approach to appraisals). On the basis of impressionistic evidence alone, in a few U.S. Schools sites where the teachers were seen as being more involved in plans for the achievement of school goals and where individual plans were seen as making a contribution to organization goals, the teachers seemed to be working more truly as professional educators.

4. CONCLUSION

A preliminary conclusion which emerges from both the literature review and the field study parts of the project is that a distinction is not consistently made between appraisal and evaluation of performance. The two terms are used interchangably in many cases but the following distinction will be used in the conclusion of this report. The term "appraisal" will be used to refer to the formative process of gathering data about performance and giving feedback, provided that the feedback is objective or non-judgemental.

On the other hand, "evaluation" will be used to refer to judgements or decisions about the relative level of performance and the giving of feedback about the judgements. A key

aspect of the difference between appraisal and evaluation is that in appraisal, it is the subordinate who has to make judgements about whether the behaviour was effective/ineffective. In evaluation, it is the supervisor who makes these judgements. If both parties share a common understanding of what constitutes effective/ineffective performance, then appraisal will be used more than evaluation and subordinates can be more autonomous in their work.

The appraisal and evaluation of performance of professionals has posed considerable difficulties for many organizations for at least three major reasons. One reason lies in the nature of professional bureaucracies where professionals perform their service work in a largely autonomous manner. As a result of their professional training (and in many cases, licensing or certification), they are assumed to be competent for the autonomous performance of their duties and assumed to be able to act in a self-regulating manner. Under these assumptions, individual accountability for professional performance is only subject to question when extreme exceptions occur to accepted norms of professional conduct. Performance which is an exception to these norms is outside the commonly shared understanding of what constitutes acceptable performance and the delegation of autonomy is revoked.

The second reason for difficulty in appraisal and evaluation of professionals is that the means and techniques (process) of a professional's work and the results (output) may be difficult to measure or assess objectively. If the work of a professional is preventative in nature, then the absence of measurable outputs may be a more valid means of appraising performance. If the results of a professional's work are qualitative or long-term in nature (i.e., to enhance feelings of self-esteem in a client; to create a sense of citizenship, and and so on), the results can be extremely difficult to assess as well. However, one should keep in mind the observation that "if the results of someone's work cannot be measured, rated, assessed or even qualitatively described in some manner to distinguish between when the work is or is not done, then we have to wonder whether any work is being done at all".[22]

The third problem lies with the assumption that a direct correspondence exists between the work of the professional and the results which are eventually obtained. In situations where numerous other variables can intervene to affect end result outcomes, this assumption can be of questionable validity. In face of the uncertainty about a performance-results correspondence, many professional bureaucracies focus appraisals on the process of professional work. If the processes are felt to be appropriate, then it is hoped that the work of the professional will be a positive contributing factor (not necessarily causal) to the desired end results.

Despite these principal difficulties, professional service organizations still have to grapple with the problem of appraisal in order to continuously try to develop or maintain a high level of professional service and to have a valid data base for personnel evaluation. This means that some form of systematic appraisal based on some means of obtaining valid data about performance has to be developed for each organization. Once the information about

---

22 Odiorne, G. "What's Wrong with Appraisal Systems?" In Personnel Policy: Issues and Practices. Columbus, Ohio: Merrill Books, 1965.

performance has been gathered, the second issue is how to make evaluations or judgements about the performance. To evaluate performance, it is necessary to have some sort of standards for purposes of comparison. Normally, evaluation requires the definition of some minimum standards of performance for at least the key aspects of a professional's work, so that the work of all persons can be judged in comparison with some fundamental requirements of the organization. Some organizations may wish to go further and define a range of standards above the minimum so that performance may be evaluated with respect to a hierarchy of standards. This will depend upon the organization's philosophy or policies about promotions and careers for professionals or the policy about performance-related compensation differentials.

4.1 OBJECTIVES AND PURPOSES OF AN APPRAISAL SYSTEM

In order for an organization to make an intelligent selection of an appropriate mix of appraisal methods, it is necessary for the organization to decide on priorities of the purposes of the appraisal system. The recommended general principle is that there should be consistency between the type of appraisal and the purpose being served by the appraisal. For example, if the organization has a very strong belief in the efficacy of a particular approach to teaching, then the appraisal system should be activities-based to ensure that the desired methods are being used. If the organization has a variety of purposes to be served by appraisal then a variety of appropriate methods will have to be used to form a multi-methods package for appraisal and evaluation of performance.

All of the organizations in the field study had some form of multi-methods appraisals. No one approach to appraisals was found to be appropriate for all purposes and each combination of methods had its strengths and limitations. Some of the multi-method appraisals were dominated by competency and activities-based appraisals, others were dominated by various forms of results-based appraisals. The range of approaches was equally varied between businesses, colleges, and schools.

While it seems self-evident that an appraisal system should be developed to serve organizational objectives and purposes, it is less self-evident which type of system serves which type of purposes. A purpose which may be common to all organizations is the need for an appraisal system to provide an objective, valid, and reliable data base for routine personnel administration and decisions. Over and above these purposes, the need is for an appraisal data base to provide information for special purposes or situations in an organization which may be different from other organizations. The following sections represent some of the special purposes which were found in the field study.

4.1.1 Situational Factors

Some of the U.S. Schools were in situations of rapid population growth where the demand for teachers exceeded the normal supply. This meant that several teachers had to be recruited from out-of-state, where little was known about different qualities of teacher training. Secondly, the quality of teacher training within the state was known to vary greatly as well.

In another case, some U.S. Schools were confronted with a crisis of a lack of confidence in the school system's ability to deliver a quality education. Since enrolment was static or even declining, it was not possible to use the approach of trying to change the quality of inputs of new teachers and the school system had to work with available resources.

A third type of situation found more in colleges and businesses was the need to develop or maintain a competitive distinction in quality of professional service. The situation faced by these organizations was that if they slipped or deteriorated in professional excellence, their clients/students could easily go to competitors who provided better service/education. A fourth situation for a number of organizations was slow or static growth with reasonably high standards already achieved. For organizations in these situations, the challenge was to both maintain the level of quality of service and provide for continued professional development of staff.

For these different organization situations, serious consideration has to be given to the question(s) of what needs or purposes are to be served by the appraisal system. For schools in high growth situations, the focus of appraisals was on controlling the standards of entry of professionals into the organization. For schools in near-crisis situations, the focus was on specific re-direction of available resources to different methods and priorities in educational services. For organizations in highly competitive situations, the focus was on excellence of service and market sensitivity for competitive advantage. For stable, good quality schools, the focus was on maintenance of standards and continual upgrading and development. These patterns are not meant to suggest that only these combinations of situations and approaches to appraisal will be suitable. Instead, the conclusion is that the focus of the appraisal system should be relevant to the situation faced by a particular school jurisdiction and that the appraisal systems may need to be reorganized from time to time to maintain a match with evolving organization situations.

4.1.2 Sense of Mission and Normative Models of Professional Service

An explicit (or strongly implicit) sense of mission in an organization or normative model of professional practice has the effect of defining a number of priorities for resource allocation and boundaries for acceptable and nonacceptable activities and behaviours. In effect, the organization creates a particular culture which then influences behaviour within the organization community in a manner analogous to the influence of a societal culture on a village. If the organization has a strong conviction that a particular method of professional service will provide the best results,[23] satisfy a purpose, or achieve a mission, then the appraisal system can be designed to elicit behaviours which fit the normative model. Normally this is done by use of an activities-based or process appraisal system, but can also be done using a results- or quasi-results-based appraisal system. The key question about a normative model is whether it has any validity for actually being an effective or more effective method of performing some professional service. Conceivably, an appraisal system could carefully

---

23 See for example, Skinner, B.F. "The Shame of American Education". American Psychologist 39, no. 9 (Sept. 1984), pp. 947-956, which advocates programmed learning.

monitor professional activities for correspondence to the normative model, but the organization goals or mission may still not be achieved if the model is unsound.

Among the organizations in this study, only a few had explicit normative models for professional services which were based on theories published by respected authors in the field. Almost none attempted any sort of validation research to see if professional service corresponding to the theory actually generated better professional service or results. Most organizations which used some sort of normative models based them on face validity, i.e., "common-sense" or collective judgements by task-forces about what constituted or was necessary for good professional service. If these collective judgements are valid, then appraisals based on these models can also be valid.

The conclusion here is that to the extent that an organization has an explicit mission or normative model for professional service, the appraisal system can be used as a powerful technique for achieving the mission or implementing the normative model. However, it is necessary for an organization to first clarify whether there is a strong sense of particular mission for which collective energies should be marshalled, or a strong conviction about a normative model of professional service to which collective energies should conform. If either of these situations is true, then the appraisal system could be designed to specifically contribute toward the achievement of these purposes.

4.1.3 Career Stages of Professionals

Most of the organizations in this study had an implicit system of different approaches to appraisal for different stages in a professional's career while an explicit distinction existed between types of appraisal for professional service and supervisory levels of the hierarchy.

The conclusion here is that what works informally and effectively in appraisals in many of the field site organizations could also work more formally as a means to reduce paperwork and increase the relevance of appraisals. To this end, an organization could use a comprehensive and multi-faceted appraisal system in recognition that some aspects will or will not be relevant to different persons, or change in relevance for individuals over time. Thus if a supervisor established that a person has satisfied performance requirements for one level of professional service work, the person is assumed thereafter to be competent in this dimension of performance unless evidence turns up to suggest the contrary.

A comprehensive, multi-faceted, and selective approach to appraisals can capture the wholeness of a professional's work and yet avoid redundancy of inappropriate appraisals. The system works best for those individuals who are satisfactory or better performers and who continue to make progress in the organization. If people fail to maintain standards or make required progress, then the more comprehensive parts of the appraisal system come into play again and, in effect, a special performance appraisal system is established for that person.

4.1.4 Absence of Mission or Normative Model

Despite the importance of having an organization mission and normative model for professional service, many organizations will not be able to specify a normative model. Instead, the organization may propose a very general mission, such as "provide good quality education" and remain non-specific on what is good quality education, or how it is to be achieved.

Two strategies for appraisal seem to be commonly used in such situations. The first strategy is to try to recruit and select the best possible candidates at the outset and turn them loose to do the best that they can. If the majority of the human resource inputs into the organization are the best possible quality, then it is assumed that they will somehow generate good results from their collective effort.

The second strategy used in the absence of explicit mission or normative models is to delegate these decisions to the professionals themselves. The strategy of delegation of decisions on missions and methods to professionals can work if the organization has a high degree of confidence in the professional component. If either or both the trustees who provide the financial resources or the clients who receive the professional service decide in some way that they are satisfied with the services, then confidence is established for delegation of self-regulating responsibility to the professionals and administrators.

If a crisis of confidence occurs, then the community, through a variety of representatives, must act to reassert responsibility and authority. This relationship of community vis-à-vis a profession must hold for virtually any profession because the community provides the resources for the profession and the ultimate justification for any profession is service to the community.

If definition of normative models of professional service is delegated to the professionals themselves, appraisals can be greatly clarified if the professionals achieve a consensus on the normative model. However, a common situation in professional bureaucracies is that little consensus exists among professionals about the most effective methods and approaches to provide professional service. Consequently, it is difficult for the professionals to be self-regulating if they have significantly divergent views about what constitutes good professional service or results. Appraisal in these circumstances reverts back to impressionistic judgements about "good" performance based on non-specific, subjective, and non-aggregated criteria.

4.2 DESIGN, IMPLEMENTATION, AND SUPPORT OF APPRAISAL SYSTEMS

Once questions have been answered about purposes, missions, and so on, of appraisal systems, the next stage, chronologically, is design and implementation of the system.

Generally there were two approaches to design and implementation used by the organizations in this study, the top-down approach and the consultative approach. In the top-down approach, the purposes, design, and implementation of a system are determined by

senior management with the aid of the personnel department and conveyed to employees as part of the conditions of employment in the organization. This approach was used by most of the businesses and about half of the colleges and schools. In the consultative approach, senior management invites the participation of subordinates in various forms of joint-committees to work on the development of a system. Usually, in the consultative approach, senior management has decided on the major purposes of a system, but invites consultation on the design and implementation, particularly for the operational details. The consultative approach was used by a few businesses and the rest of the colleges and schools.

In the consultative approach, senior management made the fundamental decisions on either basic purposes, missions, or long range goals to be served by an appraisal system, while the decision-making about operational aspects was delegated to local school boards, departments, branches, and so on, for implementation. In other words, the central authority defines a common core of appraisal aspects which have to be satisfied by all systems, but leaves definition of the other aspects up to local joint-committees for decision making. If the common core is very comprehensive and specified in great operational detail, then a top-down system is essentially being used because very little scope is left available for local decision-making. If the common core is essentially a statement of principles and philosophies with little operational definition, then a great deal of scope is delegated to local decision-making. The scope and comprehensiveness of the common core is a matter determined by the degree to which the centrality authority is willing to trust or has confidence in the local decision makers.

In terms of implementation, all of the organizations developed extensive manuals and procedures for orientation of supervisors and subordinates to the appraisal system, as well as guidance in conducting appraisals and documenting the process and the outcomes. However, most training was concerned with familiarization of supervisors and subordinates with the formal manuals and procedures so that implementation concentrated on technical knowledge. In a few cases, procedural training was supplemented by role-playing exercises where supervisors could practice the application of their technical knowledge. Training tended to be for the introductory launch only or for newly promoted supervisors. There was very little evidence of follow-up training, except on a voluntary basis in some organizations. Maintenance of the systems seemed to depend in many cases upon the personnel department's administrative procedures or demands for periodic appraisals and pressures for documentation. In some organizations, the outcome of this form of maintenance was the gradual development of the attitude that performance appraisal was done to satisfy the personnel department or some centralized bureaucratic procedures.

The conclusion is that implementation of an appraisal system was more effective when the system was seen as being either locally or jointly developed (local ownership), or a management system (management owned) rather than a personnel department program (bureaucratically owned).

4.3 METHODS AND APPROACHES TO APPRAISAL

The conclusion about approaches is that there is no one best approach and that most organizations will have to use some form of multi-faceted approach. Which combination and which emphases will be chosen will depend upon the needs, priorities, and purposes of each organization.

The disadvantage of a multi-faceted appraisal system is that principals will have to spend more time thinking about and discussing what types of appraisal are relevant for individual subordinates, and this will be a burden to principals with scarce time resources and wide spans of control. Many principals, for good reason, prefer the simplicity of a standardized checklist form of appraisal.

The advantage of a multi-faceted system is that when principals do spend more time thinking about and discussing an individual's performance and deciding what is relevant in appraisals, then the quality of individual supervision and management of human resources in a school is likely to be enhanced. Principals would have to think about different teachers as individuals, with different strengths, weaknesses, needs, and stages of development. If serious dialogue occurs between principals and teachers on these issues and if appraisals are tailored appropriately, then a considerable improvement in the quality of supervision and a more effective utilization of human resources should result. Furthermore, the problem of redundant or non-relevant appraisals should also be avoided for areas of performance where satisfactory standards have been well established.

4.4 DIMENSIONS AND CRITERIA FOR APPRAISALS

As previously stated, the foundations for the use of any set of performance dimensions for appraisal should be some theory, philosophy, or normative model of good quality professional service. Once some sort of normative model has been developed, then related performance dimensions and required competencies can be identified. The competencies should identify who can do the professional work, and the performance dimensions should identify the types of behaviours that produce desirable professional work.

The key aspect to performance dimensions and criteria is that they have to be valid representations of the performance the organization had defined as being desirable or had chosen as a normative model. If an organization truly believes in the efficacy of its particular normative model, then it should be possible to also specify some results criteria which should be observable or testable from the use of the model. If the results criteria are actually produced, then the use of the model is justified or the validity of the model is established. In practice, very few of the organizations in the field sites had a well-developed or explicit normative model of professional behaviour as a foundation for their appraisal systems, and even fewer used results criteria to verify their model.

The second challenge to development of criteria is to describe or find indicators of a range of performance for each dimension of performance. A particular criterion may consist of a two-point scale (anything below is unacceptable, anything above is acceptable) or some form of continuum (usable for development purposes or to distinguish relative levels of performance). If the criteria are highly explicit so that performance can be rated accurately and with reliability by external observers, then the criteria are said to be objective. Where criteria are not explicit for a performance dimension or cannot be rated with reliability, then the criteria are said to be subjective. While everyone seeks to develop objective criteria for performance, the reality is that many criteria are subjective and hence subject to dispute. Given this reality, the key issue seems to be acceptability of criteria more than objectivity. Where there is low acceptability about a performance dimension or how it is measured, then there is a strong need felt for objective criteria for performance. Where there is mutual confidence and acceptance between supervisors and subordinates, then subjectivity in appraisal criteria is more acceptable.

The conclusion here is that performance dimensions and various forms of criteria can only be compromises of varying degree. No list of performance dimensions can be comprehensive, and the best that can be hoped for is that the list will be reasonably representative of the important dimensions. Secondly, few criteria will satisfactorily represent the range and complexity of performance, especially for professional service types of work. Consequently, the realistic limitations of an appraisal system, no matter how complex and detailed, have to be recognized. Perhaps the most that can be expected or hoped for is a system which manages with accuracy at a level better than chance, to distinguish between acceptable (satisfies minimum standards) and less-than-acceptable performance. In addition some organizations may also wish to distinguish between performance which is acceptable and performance which is above the average.

4.5 SOURCES OF PERFORMANCE DATA FOR APPRAISALS

The review of the literature and the field study results considered four basic sources of information about the work of professionals: self-reports by professionals, client reports and results, peer reports, and supervisor monitoring and observations. All of the sources of information are valid in the sense that the work of a professional impacts on all of these components of the organization and each component is a stake-holder (i.e., affected by) in the work of the professional. However, the frequency of use of the different sources of information was quite mixed in actual practice.

Self-reports seem to be normally used for individuals with experience or seniority who are relatively well established as effective performers. The use of self-reporting is a mark of recognition for the established competence or capability of the person to be largely self-managing as a professional. Conversely, some organizations used short-term self-reporting for individuals in special performance or renewed probation situations to force the problem performer to confront the need to take some responsibility for his or her performance.

Reports and results by clients/students and subordinates were commonly used in very market-oriented businesses and colleges as an important source of information for appraisal. The use of these sources is consistent with the mission of these organizations to be market-oriented and to succeed by providing excellence of professional service.

Supervisor monitoring and appraisals, using both formal and informal information, was the most common source of data for appraisals in all of the organizations. The key role of the supervisor in the appraisal process has been strongly pointed out in a number of parts of this report. It seems to be inescapable that the ultimate appraisal of a professional's performance rests with the information gathered and judgements made by supervisors. This conclusion holds as much for an appraisal system that is highly formalistic, legalistic, and "objective" as for a system which is informal, subjective, and non-specific. Consequently it follows, and cannot be stressed too strongly, that supervisor training at all ranks in the organization is crucial to the successful functioning of an appraisal system.

Each organization has to choose which sources of information are to be used in addition to supervisor monitoring for the gathering of data for performance appraisal. The use of peer appraisals are appropriate for development purposes and where the regulation of professional services is delegated to the professionals themselves. Self-reports are appropriate for development purposes and where activities-based and results-based appraisals are used. Client reports/results are appropriate where a strong sense of mission/normative model exists or where program evaluation is desired to verify the results of a mission or model. None of these approaches or purposes are mutually exclusive and a suitable mixture will have to be chosen to satisfy the various needs and priorities of the organization.

4.6 STANDARDS FOR EVALUATION OF PERFORMANCE

As previously discussed, a distinction is made between appraisal (gathering information and providing feedback) of performance and evaluation, which concerns decision-making about the acceptability of performance. At the most basic level, evaluation is concerned with the decision about whether performance satisfies minimum acceptable standards. Normally, this type of evaluation occurs for two phases of a person's career; minimum standards to satisfy probation requirements, and minimum standards to satisfy requirements for continuity of employment or maintenance of seniority. Minimum standards are not necessarily low standards and can evolve over time as environmental conditions and organization goals change. A second type of evaluation is concerned with distinguishing between levels of acceptable performance or to distinguish between above-average and average performers in the organization.

From the results of this study, it was clear that evaluations depended upon the ability of supervisors to make judgements about performance which reduced complex appraisal information into simplistic summaries about overall performance levels. To the extent that an organization uses such evaluations for personnel decisions (compensation, training, promotions, renewed probation), this places another heavy burden of responsibility on supervisors to make complex judgements. The literature review and some of the field site results show that many supervisors try to cope with this responsibility by evaluating most

subordinates as acceptable or average. This practice has turned out to have unfortunate subsequent consequences for some organizations. The conclusion once again is that supervisor training and organizational support is crucial for effective and fair evaluations of performance.

Another caution about evaluations of performance is that they should be appropriate for the needs of the organization for management of human resources. This was a significant difference found between the business and the educational organizations. It was very important in the Business group to distinguish between levels of performance of individuals because these evaluations were the determining factor for personnel decisions about performance-based compensation, job assignments, and promotions. These sorts of human resource management decisions were far less common in the colleges and schools, although more prevalent in the colleges. Consequently, there was far less need in the educational organizations for performance evaluations to distinguish between levels of performance or to try to identify potential for future types of performance.

Minimum standards of performance for administrators may be very difficult to establish because of the variety of administrative responsibilities and the fact that so much of an administrator's efforts are devoted towards negotiated goals. Since these can change from year to year, there is less stability of content to an administrator's job than a teacher's job. One way to approach the question of evaluation for a supervisor's job is to think of it in terms of three major components:

a) routines which must be satisfactorily done year in and year out,
b) problems which are exceptions to routines or occur outside of routines and which must be dealt with as they arise,
c) changes which are to be worked on as parts of longer range goals and planning activities.

A reasonable minimum standard of performance or the minimum level of accountability to be expected of administrators should be the accomplishment of routines for which that job is responsible. Thus no matter who is the incumbent in the position, the organization has the right to expect that the routines will be satisfied.

Some individuals will have a better record at dealing with problems and exceptions than others, but the organization has the right as well to expect administrators to have some degree of competence in this aspect of management. What sort of evaluation is made of a person's performance in this area would depend upon the effort by the individual and the degree of seriousness of the problem or exception.

Finally, performance in the area of changes or goals represents either specific commitments by administrators or specific assignments to administrators. These aspects of performance represent responsibilities or accountabilities which are different from routines and which act to individualize a person's job description. Since goals and plans are specific commitments by individuals, it is appropriate for them to be held accountable for the results

(or quasi-results). The evaluation of performance in this area will depend upon the assessments made about degrees of progress towards various goals and relative importance of different goals.

4.7 SPECIAL REVIEW AND RENEWED PROBATION

The results for special performance review and renewed probation are relatively consistent for all the organizations studied. While the details and formats may differ from site to site, the general principles are the same. The personnel policies at each organization endeavour to provide for due process and fair treatment of each individual found to have a performance problem. Salvage programs are used where individuals receive feedback and counselling and are provided with opportunities to recover their former satisfactory level of performance (or build up to a required performance level if new standards are established). Progress during the period of renewed probation is carefully monitored and time limits are established to allow a fair but finite opportunity to recover. If the individual apparently cannot satisfy the minimum required standards despite repeated trials and salvage programs, then the individual has demonstrated lack of competency to perform the required type or level of professional service.

Differences between the study sites were in three areas of practice. Some organizations had established more explicit minimum standards of performance than others with the effect that individuals could be more self-monitoring and recognize when their performance was near a danger zone. Some organizations allowed longer time periods for renewed probation or salvage programs with the effect that the individual had a longer time available to recover or re-establish necessary performance levels (or a longer time to avoid doing anything). The organizations which tended to be more explicit and demanding in performance standards and quicker to take action on performance problems, were those who tended to be more market-sensitive or with firm policies of professional service excellence.

The conclusion here is that if the organization has a firm commitment to a strategy of market-sensitivity or a firm commitment to excellence of its professional services, then a number of policies and practices in supervision, operational procedures, and personnel policies have to be consistent with the goals or act to give force to these goals. The approaches taken by an organization in requiring high standards of performance, being explicit about standards, and taking care to maintain standards are an example of this need for consistency. An organization which operates in a quasi-monopoly situation or enjoys secure access to resources might not feel such a need for market sensitivity or accountability for use of resources and, hence, could afford to be less explicit and demanding about levels of professional service offered to the public.

4.8 PROGRAM EVALUATION AND MAINTENANCE

In section 4.1 it was stressed that an organization should try to define goals and purposes to be served by a performance appraisal program. Secondly, it was stressed that goals and purposes should be related to the special needs or situation faced by an organization in

relation to its environment. The subsequent topics of design, methods, and implementation were all concerned with the creation of an organization mechanism (the appraisal system) for the achievement of, or the direction of effort towards, those goals and purposes. Program evaluation is concerned with completing the cycle by examining whether the appraisal mechanism does in fact work towards or achieve those goals.

While it would seem that program evaluation is a crucial activity in the management of an appraisal system, and is consistently recommended in the literature, it was actually found rarely in practice. The majority of the appraisal programs were informally evaluated by managers' subjective and impressionistic evidence about whether a program was "working well" or relatively free of problems. If programs were felt to be satisfactory as a result of such impressions, then it was assumed that progress was being made toward the achievement of organization goals.

This can be a risky approach to evaluation of a program because satisfactory progress on established procedures and lack of dis-confirming short-term evidence are only suggestive of or consistent with achievement of goals and purposes, but cannot be equated with such achievement. Consequently it is of great importance to follow through with program evaluation to ensure that the organization's resources are not being wasted or that the appraisal system is actually fulfilling its role. The fact that so few of the organizations were using a formal approach to program evaluation suggests that either they do not realize the risk of such an omission or that they have great confidence in the appropriateness of their particular approach to appraisals.

4.9 SUMMARY CONCLUSIONS

In summary, three general conclusions are derived from the results of this study.

a) Different approaches and priorities in appraisal systems are used for relevance to the local situation of a particular organization or school jurisdiction. This requires all participants in the development of a system to carefully consider the needs and characteristics of their situation and carefully develop goals and purposes appropriate for their situation.

b) Appraisal systems can be expected to go through stages of evolution as all participants have to go through a learning curve of experience with a system. Therefore no one design can be expected to be correct at the outset nor remain in force indefinitely. Different approaches to appraisal may have to be used for different stages of individual career cycles, different levels of jobs in the hierarchy, and for different levels of experience with a system.

c) An appraisal system developed jointly between supervisors and subordinates has a better chance of incorporating diverse but relevant points of view than a system developed centrally by top management alone. The greater the opportunity for participation by parties affected by a decision, the greater the potential quality and acceptability of the decisions. However, a need also exists in large organizations to ensure that equity, fairness, and minimum standards are provided for across all systems of appraisal. A workable solution seems to be a centrally developed (with joint participation) system of universal standards which

have to be satisfied by any appraisal system, and then delegation to local authorities (or joint committees) for the development of local, operational specifics.

In the end, no system of appraisals can be workable, no matter how carefully designed, if the people who have to work with it and live by it do not view it as being acceptable and legitimate. Representative participation in the development of a system was not found to be absolutely necessary for acceptability and legitimacy of a system, but participation was found to greatly enhance acceptability. The other side of the coin is that the participants in the joint committee have to share in the goals of an organization or share in the commitment to professional service. If various participants in the planning process use participation as an opportunity for the pursuit of other political goals, then the necessary mutuality of goals will not be present and decision quality and acceptability will not be enhanced. This caveat aside, the probability of developing an effective and appropriate appraisal system is greatly improved through the use of joint planning and development of the system.

APPENDIX 1

CONDUCTING APPRAISAL INTERVIEWS

Step 1: Prepare

- The employee should be contacted several days in advance to allow time to prepare.
- It should be a mutually convenient time.
- Assemble all data (documentation, appraisal instruments, and so on) and refamiliarize yourself with them.
- Ensure that there will be privacy, lack of interruption, sufficient time.

Step 2: Outline the Purpose and Set the Tone

- Make subordinate comfortable.
- Get to the point of the interview (e.g., "As we discussed when we arranged this meeting, our purpose today is to discuss your performance since your last review and what we might do together to help you with any problems you may be having").
- Describe how you intend to conduct the meeting.
- Emphasize two-way communication.

Step 3: Invite Employee's Assessment of His/Her Performance

- Start the discussion by asking "How do you feel things are going on the job? What's going well and what problems are you having? How do you see your performance?" and so on
- If the employee had performance objectives, have your subordinate restate each objective and results obtained.
- Ask questions for purposes of clarification and elaboration.

Step 4: Listen and Learn

- Do not interrupt.
- Listen to what the employee is saying.
- Be prepared to alter your assessment in light of new data.

Step 5: Share Your Data

- Give specific examples of proficiencies and deficiencies.
- Emphasize employee strengths.
- Discuss performance, not personality.
- Tell employees what is wrong with their work, not themselves.
- Avoid comparing the employee to others.
- Allow the employee the opportunity to ask questions, clarify information.
- Explain why you assessed performance as you did.
- When discussing weak areas, focus on solving the problem (causes and possible remedies) rather than on the event.

Step 6: Summarize Your Joint Analysis

- Discuss various aspects of your data and the data of the employee; encourage mutual feedback.
- Praise the positive aspects of performance.
- Review what has been stated, and on what points you agree.
- Discuss where you differ, why you differ (remember complete agreement is not always mandatory), and how to resolve differences relating to future job performance.

Step 7: Develop a Plan of Action

- Emphasize future improvement in performance.
- Encourage the employee to think out solutions to problems (e.g., What do you feel is the problem here? What do you think can be done? How? How can I help?).
- Suggest specific techniques to help improve performance.
- Set formal or informal performance targets and goals (process- or output-oriented).
- Outline jointly what changes are to be expected in the employee's performance in specific terms.
- Review action plans and set follow-up dates to review progress.
- End interview on a positive note of encouragement.

Step 8: Follow-Up

- Summarize the meeting in writing.
- Ensure employee has a copy of appraisal.
- Meet with subordinate on agreed upon date to review progress.
- Make yourself available for coaching and counselling.
- Monitor employee for signs of improvement and give positive feedback.

APPENDIX 2

QUESTIONS ABOUT THE RELATIONSHIP BETWEEN ADMINISTRATOR AND TEACHER APPRAISAL

(1) What proportion of the job functions of teachers and administrators are interactive rather than mutually exclusive?

(2) Should teachers appraise principals?

(3) Should effective teacher appraisal be an objective of administrator appraisal?

(4) How is educational leadership tied to performance evaluation?

(5) To what extent should administrators be evaluated on their ability to evaluate teachers?

(6) Are organizational goals linked to performance at all levels of the school hierarchy?

(7) Should the philosophy and overall objectives of appraisal be the same for both teachers and administrators?

(8) Is it feasible to use the same fundamental process (e.g., number of appraisers, assessment period, data sources, instrument design, goal-setting processes, competency-based measurement) for all school-based personnel?

(9) Do differences in tasks, responsibilities, and criteria for assessment between organizational levels impede the design of a congruent system of appraisal?

(10) Would implementation of the same appraisal process at all levels within the school streamline evaluation? facilitate communication? reduce resistance to the process? serve as a co-ordinating device among levels?

REFERENCES

Acheson, K.A., and Gall, M.D. Techniques in the Clinical Supervision of Teachers. New York: Longman, 1980.

Alexander, E.R., and Wilkins, R. D. "Performance Rating Validity: The Relationship of Objective and Subjective Measures of Performance". Group and Organization Studies 7, no. 4 (December 1982), pp. 485-96.

Appel, J. "Yearly Evaluation - A Key in Assessing Superintendents". Thrust For Educational Leadership (January 1980), p. 28.

Bauman, K.E., and Dent C.W. "Influence of an Objective Measure on Self-Reports of Behavior". Journal of Applied Psychology 67, no. 5 (October 1982), pp. 623-628.

Bazerman, M.H.; Rafik, I.B.; and Schoorman, F.D. "Performance Evaluation in a Dynamic Context: A Laboratory Study of the Impact of a Prior Commitment to the Ratee". Journal of Applied Psychology 67 (December 1982), pp. 873-876.

Bernardin, H.J., and Cardy, R.L. "Appraisal Accuracy: The Ability and Motivation to Remember the Past". Public Personnel Management 11, no. 4 (Winter 1982), pp. 352-357.

Bernardin, H.J., and Beatty, R.W. Performance Appraisal: Assessing Human Behavior at Work. Boston: Kent, 1984.

Blai, B. Jr. "An Appraisal System That Yields Results". Supervisory Management 28 (November 1983), pp. 39-42.

Brush, D.H., and Schoenfeldt, L.F. "Performance Appraisal for the '80s". Personnel Administrator 27 (December 1982), pp. 76-83.

Buser, R.L., and Banks, F.A. Jr. "The Why, What, How and By Whom of Evaluating Principals". NASSP 68 (January 1984), pp. 1-4.

"Can the Schools Be Saved?" Newsweek, May 9, 1983.

Carroll, S.J. Jr., and Schneier, C.E. Performance Appraisal and Review Systems: The Identification, Measurement, and Development of Performance in Organizations. Glenview, Ill: Scott Foresman, 1982.

Caruth, D.; Middlebrook, B.; and Rachel, F. "Performance Appraisals: Much More Than a Once-a-Year Task". Supervisory Management 27, no. 9 (September 1982), pp. 28-36.

Cascio, W.F. "Scientific, Legal, and Operational Imperatives of Workable Performance Appraisal Systems". Public Personnel Management 11, no. 4 (Winter 1982), pp. 367-375.

Cederblom, D. "The Performance Appraisal Interview: A Review, Implications and Suggestions". Academy of Management Review 7, no. 2 (April 1982), pp. 219-227.

Cohen, S.L., and Jaffee, C.L. "Managing Human Performance for Productivity". Training and Development Joural 36 (December 1982), pp. 94-95.

Cordes, C. "Research Finds Little Merit in Merit Pay". Monitor 14, no. 9 (September 1983), p. 1.

Darling-Hammond, L.; Wise, A.E.; and Pease, S.R. "Teacher Evaluation in the Organizational Context: A Review of the Literature". Review of Educational Research 53, no. 3 (Fall 1983), pp. 285-328.

De Bevoise, W. "Synthesis of Research on the Principal as Instructional Leader". Educational Leadership 41, no. 5 (February 1984), pp. 15-20.

Deal, T.E.; Dornbusch, S.M.; and Crawford, R.A. "Villians as Victims: Evaluating Principals". Phi Delta Kappan (December 1979), pp. 273-74.

DeMarco, J.J., and Nigro, L.G. "Implementing Performance Appraisal Reform in the United States Civil Service". Public Administration 61 (Spring 1983), pp. 45-57.

DeNisi, A.S.; Randolph, W.A.; and Blencoe, A.G. "Potential Problems With Peer Ratings". Academy of Management Journal 26 (September 1983), pp. 457-464.

Dittloff, R. "Evaluate the Superintendent". American School Board Journal 169, no. 11 (November 1982), p. 11.

Duhamel, R.; Cyze, M.; Lamacraft, G.; and Rutherford, C. "The Evaluation of Principals" Education Canada 21, no. 2 (Summer 1981), pp. 20-27.

Edwards, M.R. "Productivity Improvement Through Innovations in Performance Appraisal". Public Personnel Management 12 (Spring 1983), pp. 13-24.

Edwards, M.R., and Sproull, J.R. "Rating the Raters Improves Performance Appraisals". Personnel Administrator 28 (August 1983), pp. 77-82.

Farquhar, R.H. "How the Teaching Profession Measures Teaching Effectiveness". Paper presented at the conference on "Teaching in Medical School," Saskatoon, Saskatchewan, June 1978. (ED 1985 683).

Feldvebel, A.M. "Teacher Evaluation: Ingredients of a Credible Model". Clearing House (May 1980), pp. 415-420.

Fisher, C.D., and Thomas, J. "The Other Face of Performance Appraisal". Human Resource Management (Spring 1982), pp. 24-26.

Fleenor, C.P., and Scontrino, M.P. Performance Appraisal: A Manager's Guide. Dubuque, Iowa: Kendall/Hunt, 1982.

Flynn, J.J. "Evaluation of School Administrators". Comment on Education (October 1978), pp. 5-8.

Fordham, A. "Evaluation Reflection: The Context of Teaching and Learning in Studies of Teacher Effectiveness". Studies in Educational Evaluation 8, no. 2 (1982), pp. 111-128.

Gallegos, P.M. "Communicating Performance Results". Journal of Systems Management 34, no. 3 (March 1983), pp. 25-31.

Gelberd, L.B.; Lee, C.; and Lupton, D.E. "Self-Appraisals: The Next Stage in Performance Evaluations". Supervisory Management 28 (May 1983), pp. 9-14.

Glueck, W.F. Personnel: A Diagnostic Approach. Plant, Texas: Business Publications, Inc. 1982.

Gomez-Mejia, L.R, and Page, R.C. "Integrating Employee Development and Performance Appraisal". Training and Development Journal 37 (June 1983), pp. 138-145.

Graves, J.P. "Let's Put Apraisal Back in Performance Appraisal: Part 1". Personnel Journal 61 (November 1982), pp. 844-849.

Graves, J.P. "Let's Put Appraisal Back in Performance Appraisal: Part 2". Personnel Journal 61 (December 1982), pp. 918-923.

Guba, E.G., and Lincoln, Y.S. Effective Evaluation. San Fransisco: Jossey-Bass, Inc, 1981.

Gutridge, B.M. Teacher Competency: Problems and Solutions. American Association of School Administrators, 1980.

Hathaway, W.E. "Testing Teachers". Educational Leadership 33 (1980), pp. 210-215.

Hautaluoma, J.E.; Kaman, V.S.; and Kryger, B. "Performance Appraisal For Human Resource Management: Implications of a Pre-Post Study". Human Resource Planning 6, no. 1 (1983), pp. 41-48.

Herman, J.J. "An Up-Front Approach to Administrative Evaluation". NASSP Bulletin. (November 1978), pp. 34-42.

Hewitt, P.J. "Performance Evaluation: How To Make It Easier on You - and Your Employee". Supervision 45 (June 1983), pp. 14-15.

Holt, M. Evaluating the Evaluators. London: Hodder and Stoughton, 1981.

Hom, P.W.; DeNisi, A.S.; Kinicki, A.J.; and Bannister, B.D. "Effectiveness of Performance Feedback From Behaviourally Anchored Rating Scales". Journal of Applied Psychology 67, no. 5 (October 1982), pp. 568-576.

Huber, V.L. "An Analysis of Performance Appraisal Practices in the Public Sector: A Review and Recommendations". Public Personnel Management 12 (Fall 1983), pp. 258-267.

Hyde, A.C., and Smith, M.A. "Performance Appraisal and Training". Public Personnel Management 11, no. 4 (Winter 1982), pp. 358-366.

Ivancevich, J.M. "Subordinates' Reactions to Performance Appraisal Interviews: A Test of Feedback and Goal-Setting Techniques". Journal of Applied Psychology 65, no. 5 (October 1982), pp. 581-587.

Iwanicki, E.F. "Development and Validation of the Teacher Evaluation Needs Identification Survey". Educational and Psychological Measurement 42, no. 1 (Spring 1982), pp. 265-74.

Kaye, B., and Krantz, S. "Performance Appraisal: A Win/Win Approach". Training and Development Journal 37, no. 3 (March 1983), pp. 32-34.

_____. "Preparing Employees: The Missing Link in Performance Appraisal Training". Personnel 59, no. 3 (May/June 1982), pp. 23-29.

Kirkpatrick, D.L. How To Improve Performance Through Appraisal and Coaching. New York: AMACOM, 1982.

Landy, F.; Zedeck, S.; and Cleveland, J. Performance Measurement and Theory. Hillsdale, New Jersey: Lawrence Erlbaum, 1983.

Lewis, A.C. Evaluating Educational Personnel. American Association of School Administrators, 1982.

Lockwood, M., and Ludmer, A. "The Performance Model: An Objective Approach to Measuring White Collar Productivity". Manage 34, no. 4 (October 1982), pp. 30-32.

Lotto, L.S. "Revisiting the Role of Organizational Effectiveness in Education Evaluation". Educational Evaluation and Policy Analysis 5, no. 3 (Fall 1983), pp. 367-378.

Manaktala, V.K. "Job Descriptions: Making Subordinates Accountable". Supervisory Management 28, no. 1 (January 1983), pp. 27-29.

McCleary, L. "Evaluation of Principals". Theory Into Practice (February 1979), pp. 45-49.

McGreal, T.L. "Developing a Teacher Evaluation System: Commonalities of Those Systems that Function Most Effectively". Paper presented at the Annual Meeting of the Association For Supervision and Curriculum Development, March 1982. (ED 226 418).

McGreal, T.L. "Effective Teacher Evaluation Systems". Educational Leadership 39, no. 4 (January 1982), pp. 303-305.

Meyer, J.W., and Rowan, B. "The Structure of Educational Organizations". In Meyer, M.M., ed. Environments and Organizations. San Francisco: Jossey-Bass, 1978.

Mintzberg, H. The Structuring of Organizations. New Jersey: Englewood Cliffs, Prentice-Hall, 1979.

Morrisey, G.L. Performance Appraisals in the Public Sector: Key to Effective Supervision. Reading, Mass: Addison-Wesley, 1983.

Mount, M.K. "Comparisons of Managerial and Employee Satisfaction With a Performance Appraisal System". Personnel Psychology 36 (Spring 1983), pp. 99-110.

Murphy, J.; Hallinger, P.; and Mitman, A. "Problems with Research on Educational Leadership: Issues To Be Addressed". Educational Evaluation and Policy Analysis 5, no. 3 (Fall 1983), pp. 297-305.

Nevo, D. "The Conceptualization of Educational Evaluation: An Analytical Review of the Literature". Review of Educational Research 53, no. 1 (Spring 1983), pp. 117-128.

Patten, T.H. A Manager's Guide to Performance Appraisal: Pride, Prejudice, and The Law of Equal Opportunity. New York: Free Press, 1982.

Peterson, D. "Legal and Ethical Issues of Teacher Evaluation: A Research Based Approach". Educational Research Quarterly 7, no. 4 (Winter 1983), pp. 6-16.

Peterdon, D., and Ward, A. eds. Due Process in Teacher Evaluation. Washington, D.C.: University Press of America, 1980.

Peterson, K., and Kauchak, D. "Progress on Development of Lines of Evidence for the Evaluation of Public School Teaching". Paper presented at the American Educational Research Association, Montreal, April 1983 (ED 228 228).

Petrie, T.A. "Ideas That Hinder Evaluation - Debunking the Myths". NASSP Bulletin 66 (December 1982), pp. 52-55.

Phelan, W.T. "Governing Staff Reductions: The Use and Abuse of Teacher Evaluations". Education and Urban Society 15, no. 2 (February 1983), pp. 189-198.

Raths, J., and Preskill, H. "Research Synthesis on Summative Evaluation of Teaching". Educational Leadership 39, no. 4 (January 1982), pp. 310-13.

Redfern, G.B. Evaluating Teachers and Administrators: A Performance Objectives Approach. Boulder, Colorado: Westview Press, 1980.

Renihan, F.I. "The Evaluation of Principals". Saskatchewan Education Administrator (December 1980), pp. 22-31.

Riley, B.E. "Essential Phases of School Personnel Management: Communication, Conferencing, Observation and Evaluation", 1982 (ED 221 543).

Rist, M.C. "Our nationwide poll: most teachers endorse the merit pay concept". American School Board Journal 170, no. 9 (September 1983), pp. 26-33.

Robinson, T.B. Jr. "A New Approach to Increasing Organizational Commitment and Productivity - The Management Planning and Performance Appraisal System". Training and Development Journal 37, no. 1 (January 1983), pp. 58-60.

Sanzotta, D. "Assessing Your Appraisal Style". Supervisory Management 27 (December 1982), pp. 2-7.

Sapone, C.V. "Appraisal and Evaluation Systems: Perceptions of Administrators, Teachers". NASSP Bulletin 65 (1981), pp. 25-30.

Sashkin, M. Assessing Performance Appraisal. Burlington, Ontario: University Associations, 1983.

Savage, J.G. "Teacher Evaluation Without Classroom Observation". NASSP Bulletin 66 (December 1982), pp. 41-45.

Schimel, R.M. "Making Performance Evaluation Work Better: A Supervisor's Viewpoint". Public Personnel Management 11, no. 4 (Winter 1982), pp. 335-337.

Schlechty, P.C., and Vance, V.S. "Recruitment, Selection, and Retention: The Shape of the Teaching Force". The Elementary School Journal 83, no. 4 (1983), pp. 469-487.

Scontrino, M.P., and Fleenor, C.P. "Performance Appraisal and Productivity: A Natural Link". Journal of Contemporary Business 11, no. 2 (1982), pp. 832.

Seidman, W.H. "Goal Ambiguity and Organizational Decoupling: The Failure of Rational Systems Program Implementation". Educational Evaluation and Policy Analysis 5, no. 4 (Winter 1983), pp. 399-413.

Silverman, B.R.S. "Why the Merit Pay System Failed in the Federal Government". Personnel Journal 62, no. 4 (April 1983), pp. 294-297.

Smith, A., ed. Evaluating Faculty and Staff. San Fransisco: Jossey-Bass, Inc, 1983.

Solomon, G. "Teachers Rating Principals: A Sobering Experience". Principal 62, no. 4 (March 1983), pp. 14-17.

Starling, G. "Performance Appraisal in the Z Organization". Public Personnel Management 11, no. 4 (Winter 1982), pp. 343-351.

Stengel, R. "Quality Not Just Quantity: The Paideia Proposal Aims to Reform America's Schools". Time, September 6, 1982, p. 87.

Stow, S.B., and Manatt, R.P. "Administrator Evaluation Tailored to Your District or Independent School". Educational Leadership 39, no. 5 (February 1982), pp. 353-357.

Strike, K.A., and Millman, J. "Nontechnical Questions About Teachers Evaluation Systems in Elementary and Secondary Schools: A Research Agenda". Educational Evaluation and Policy Analysis 5, no. 4 (Winter 1983), pp. 389-397.

Sweeney, J. "Research Synthesis on Effective School Leadership". Educational Leadership 39, no. 5 (February 1982), pp. 346-52.

Sweeney, J. "The Post-Observation Conference: Key to Teacher Improvement". The High School Journal 66, no. 2 (December-January 1983), pp. 135-140.

"Teacher Evaluation: A CTF Conference Report". Report of a Workshop Sponsored by the Canadian Teachers' Federation, September 1982. (ED 224 105).

Thompson, F.J. "The Performance Appraisal of Public Managers: Inspiration, Consensual Tests and the Margins". Public Personnel Management 11, no. 4 (Winter 1982), pp. 306-313.

Toll, S. "This merit pay plan for management improves performance and sets an example for other employees (like teachers, for instance)". American School Board Journal 170, no. 8 (August 1983), pp. 38-9.

Totusek, P.F., and Christner, C.A. "So You Think You Want to Develop an Administrator Evaluation System". Paper presented at the Annual Meeting of the Amerjcan Educational Research Association, April 1983. (ED 228 294).

Walters, D.L. "The Measurement of Administrative Competencies". Phi Delta Kappan (February 1980), pp. 423-425.

Weick, K.E. "Administered Education in Loosely Coupled Schools". Phi Delta Kappan 63, no. 10 (1982), pp. 673-76.

Weick, K.E. "Educational Organizations as Loosely Coupled Systems". Administrative Science Quarterly 21 (March 1976), pp. 1-19.

Wells, R.G. "Guidelines For Effective and Defensible Performance Appraisal Systems". Personnel Journal (October 1982), pp. 776-782.

Whitbeck, P.H. "A Critique on the Theme of Change in Performance Appraisal". Public Personnel Management 11, no. 4 (Winter 1982), pp. 340-342.

Wuhs, S.K., and Manatt, R.P. "The pace of mandated teacher evaluation picks up". American School Board Journal 170, no. 5 (May 1983), p. 28.

Zammuto, R.F.; London, M.; and Rowland, K.M. "Organization and Rater Differences in Performance Appraisals". Personnel Psychology 35, no. 3 (Autumn 1982), pp. 643-658.

Zedeck, S., and Cascio, W.F. "Performance Appraisal Decisions as a Function of Training and Purpose of the Appraisal". Journal of Applied Psychology 67 (December 1982), pp. 752-758.